AD_____

AWARD NUMBER: DAMD17-03-1-0475

TITLE: Integrating Geographic Information Systems (GIS) into Breast Cancer Epidemiologic Research

PRINCIPAL INVESTIGATOR: Daikwon Han, Ph.D.

CONTRACTING ORGANIZATION: State University of New York
Albany, New York 12201-0009

REPORT DATE: February 2006

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-02-2006	Annual Summary	1 Jul 2003 – 15 Jan 2006

4. TITLE AND SUBTITLE

Integrating Geographic Information Systems (GIS) into Breast Cancer Epidemiologic Research

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0475

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Daikwon Han, Ph.D.

E-Mail: dhan@buffalo.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

State University of New York
Albany, New York 12201-0009

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The objective of this postdoctoral training research is an integration of GIS and a spatio-temporal perspective into breast cancer research of the relationship between environmental exposures and breast cancer risk. We have completed collection of historic traffic information, and updating of the lifetime residential histories for our dataset for the breast cancer cases and controls. I have also been involved in analysis and writing of a classic epidemiologic research paper and have participated in several workshops as a part of epidemiology training. The development of a theoretical framework measuring similarity and difference of individual's lifetime residential history was completed, and a GIS-based traffic model was established. These data and epidemiologic evidence were used for further analyses, based on the models of geospatial lifeline for the estimation of lifetime residential exposures to PAHs and breast cancer risk. Findings from these studies indicate that environmental exposures in early life may be associated with breast cancer risk. Since we found evidence of association between PAHs exposures in relation to breast cancer risk, especially PAH exposures during sensitive time periods in early life, we strongly believe that models of geospatial lifeline can be effectively used for the estimation of lifetime residential exposures to PAHs and breast cancer risk.

15. SUBJECT TERMS
breast cancer, Geographic Information Systems (GIS), Environmental epidemiology

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	71	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

INTRODUCTION

The objective of this postdoctoral training research is an integration of Geographic Information Systems (GIS) and a spatio-temporal perspective into breast cancer research of the relationship between environmental exposures and breast cancer risk. With the increased use of GIS in epidemiologic studies, it becomes possible to examine lifetime exposures to environmental risk factors by integrating lifetime exposure information in GIS with other breast cancer epidemiologic factors. As a part of the ongoing case-control study of breast cancer in western New York, the proposed study examines the relationships between the residential environment and breast cancer risk and test the hypothesis, 1) Lifetime cumulative exposure to Polycyclic Aromatic Hydrocarbons (PAHs) and benzene will be more strongly associated with risk for breast cancer than of any one time period. 2) There will be sensitive periods in a woman's life that will carry greater risk for exposure. Specific aims of the proposed study include, 1) Continuing evaluation on the role of environmental risk factors on breast cancer; we will develop a GIS-based model of lifetime residential history and environmental exposure in breast cancer. 2) Assessment of historical exposures to PAHs and benzene and breast cancer risk; we will reconstruct historical exposure to PAH and benzene for the use in environmental epidemiology of breast cancer and empirically assess exposure to these two environmental compounds based on lifetime exposure index. Using lifetime residential information for breast cancer cases and controls in western New York, the proposed study examines breast cancer risk from lifetime exposures. Lifetime residential histories have been collected for the breast cancer cases and controls by interview, while information on environmental contaminants is being collected from historical sources. We use GIS in our assessment of the associations of environmental risk factors and breast cancer incidence. Further spatial-statistical analyses are performed in a GIS environment to examine associations between residential environment and breast cancer risk in spatial and temporal dimensions. Preliminary findings are discussed in the text of this report, and these findings indicate that environmental exposures in early life may be associated with breast cancer risk.

BODY OF REPORT

Task 1: Developing plan for modeling, data needs, and training.

As proposed, developing a plan for modeling and development of the data were a major component of the first year goals. For this goal, I have been focusing on completion of spatial clustering analyses of residences. I completed a GIS-based spatial and temporal analysis for residences of breast cancer cases and controls at early life and found strong evidence of spatial clustering for cases during this time. A paper on geographic clustering of residence in early life and subsequent risk of breast cancer has been published (Han, D. Rogerson, PA. Nie, J. Bonner, MR. Vena, JE. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. 2004. Geographic Clustering of Residence in Early Life and Subsequent Risk of Breast Cancer. Cancer Causes and Controls 2004;15: 921-929). My paper based on this work was selected as one of ten finalists for the Nystrom competition of the Association of American Geographers (AAG), and the paper was presented at the centennial meeting of the AAG, Philadelphia, PA. A copy of the manuscript is included in Appendix.

We also completed updating of the lifetime residential histories for our dataset for the breast cancer cases and controls. All Erie and Niagara county residential location were identified and geocoded, and these were merged into one database. We checked consistency of geocoded addresses in different time points for each individual, updated incomplete addresses using Polk searches, and validated the consistency of reported years of moved in and out of the residence. A paper regarding the clustering of lifetime residence and breast cancer risk using exploratory spatial analysis tools based on these lifetime residential history data has been accepted for publication (Han, D. Rogerson, PA. Bonner, MR. Nie, J.Vena, JE. Muti, P. Trevisan, M. Freudenheim, JL. 2005. Assessing Spatio-Temporal Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer. International Journal of Health Geographics 4:9), and an abstract based on this work was presented at the annual meeting of the International Society for Environmental Epidemiology in New York City, New York, August, 2004. A copy of the abstract is in Reportable Outcome Section, and a copy of the manuscript is included in the Appendix.

Second, we completed collection of the PAHs and benzene exposure from historic traffic information and air pollution sources. Working with my colleagues Drs. Jing Nie and Matthew Bonner and others, we have assessed historical exposure to PAH from these sources, and found evidence of association between PAHs exposures from traffic and air pollution sources in relation to breast cancer risk, especially PAH exposures during sensitive time periods in early life. Also a GIS-based traffic model was established to estimate historical residential exposure to PAHs from traffic, and a geostatistical method was utilized to predict and interpolate individual residential TSP concentration for the estimation of PAHs exposure from air pollution source. These data and evidence from epidemiologic studies will be used in my future spatial analyses of lifetime exposure to PAHs and breast cancer risk. I was a contributor to a paper regarding historical exposure to PAHs as measured by total suspended particulates and by traffic PAHs.

Training in epidemiology was another major task. I have been involved in analysis and writing of a classic epidemiologic research paper and have participated in

several workshops on epidemiologic methods. As a part of my training, I took an epidemiologic methods course "Advanced epidemiologic methods" in the Department of Social and Preventive Medicine in the University at Buffalo. I have also participated in the National Cancer Institute's summer curriculum in cancer prevention program, and took the course "Principles and practice of cancer prevention and control." I also took a cancer oncology course, "Oncology for scientists (RPN530)" at the Roswell Park Cancer Institute, Buffalo, NY in the fall of 2004. In addition, I participated in the Center for Spatially Integrated Social Science (CSISS) summer workshop, "Geographically weighted regression and associated statistics", and I was selected as one of the participants in the Harvard School of Public Health workshop, "The Public Health Disparities Geocoding Project" which I attended in June 2005.

Finally, in order to increase my skills in epidemiologic analysis, I have been working on a more classic exposure-disease analysis of lifetime body weight and breast cancer risk. The paper is however a departure from a completely classical analysis in that I am looking at early exposures and modeling cumulative exposures as well as exposures in potentially sensitive time periods; this work then is also consistent with the GIS-based work that I have been doing on early exposures and breast cancer risk. I have been using the same breast cancer study as for my GIS research, a case-control study in western New York, the Western New York Exposures and Breast Cancer (WEB) study. An abstract based on this work was presented at the annual meeting of the Society for Epidemiologic Research in Salt Lake City, June 2004. A manuscript has been submitted regarding effects of lifetime adult weight change on pre- and post-menopausal breast cancer risk using breast cancer case-control data in western New York. I have attached a copy of the manuscript submitted in the appendix. (Han, D. Nie, J. Bonner, MR. McCann, SE. Muti, P. Trevisan, M. Ramirez, F. Vito, D. Freudenheim, JL. Lifetime Adult Weight Gain, Central Adiposity, and the Risk of Pre- and Postmenopausal Breast Cancer in the Western New York Exposure and Breast Cancer (WEB) Study. Submitted).

Task 2: Developing and testing a model for lifetime exposure and breast cancer risk

The development of a theoretical framework measuring similarity and difference of individual's lifetime residential history is completed. Investigators in the Department of Geography and NCGIA have developed GIS-based theoretical frameworks measuring similarity or dissimilarity of individual's lifetime residential history in space and time (Sinha, G. and Mark, D. Measuring similarity between geospatial lifelines in studies of environmental health. Journal of Geographical Systems 2005 7:115-136). I have applied these models of geospatial lifeline to breast cancer case and control data in Western New York, and I have been focusing on the development of algorithms for missing residential history (about 7%). Further, we have considered three additional factors in developing a model for lifetime exposure and breast cancer risk. First, accuracy of historic residential information was validated. As part of this work, we have obtained birth certificates for study participants including birth address information, and have validated the accuracy of reported birth addresses relative to the birth certificates. Second, interactions of socio-economic factors with environmental exposures and subsequent risk of breast cancer were evaluated. Toward this goal, I participated in the Harvard School of Public Health geocoding workshop. I have explored the possibility of getting historic

socio-economic variables, such as area-based measures of income or education, from available historic census data. Socio-economic factors may be directly related to differential environmental exposures, and geographic clustering of lifetime residence may occur because of such socio-economic factors. Thus, we developed a model of lifetime exposure and breast cancer risk after adjusting for socio-economic variables. Lastly, variations due to genetic or biological difference were included in the model. For this goal, we have examined clustering of early life residence and breast cancer risk by genetic and biological characteristics. We found evidence of clustering among cases with a GSTM1 null genotype, and observed a tendency to clustering among those with estrogen receptor positive (ER+) tumors relative to controls. These findings indicate that there are effects of common exogenous exposures related to residence in early life, especially among subgroups of cases, and that this genetic susceptibility to environmental exposure may play a role in subsequent risk of ER+ breast cancer. A copy of the abstract is in Reportable Outcome Section.

Task 3: Assessing historical exposures to PAHs and benzene and breast cancer risk.

We have collected and analyzed the PAHs and benzene exposure information from lifetime smoking histories (active and passive smoking histories), traffic roadways, air pollution sources (total suspended particulates), and industrial sites. Also a GIS-based traffic model was established to estimate historical residential exposure to PAHs from traffic, and a geostatistical method was utilized to predict and interpolate individual residential TSP concentration for the estimation of PAHs exposure from air pollution source. As described earlier, we have assessed historical exposure to PAH from these sources, and found evidence of association between PAHs exposures in relation to breast cancer risk, especially PAH exposures during sensitive time periods in early life. Working with my colleague Dr. Matthew Bonner, Jing Nie, and others, two publications were produced; 1) Bonner, MR. Nie, J. Han, D. et al. Environmental tobacco smoke exposure in early life and the risk of breast cancer. Cancer Causes and Control 2005 16:683-689). 2) Bonner, MR. Han, D. Nie, J. et al. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. Cancer Epidemiology, Biomarkers and Prevention 2005 14:53-60. We also found evidence of association between traffic PAHs exposures in early life and breast cancer risk, and two abstracts were published; 1) Nie, J. Bonner, MR. Han, D. et al. "Environmental exposure to traffic polycyclic aromatic hydrocarbons (PAHs) and risk of breast cancer" Proc. Amer Assoc Cancer Res 2005 46: 4916. 2) Nie, J. Bonner, MR. Han, D. et al. "Traffic Polycyclic Aromatic Hydrocarbons Genetic Susceptibility and Risk of Breast Cancer" Pediatric Research 2005 58: 1025. A manuscript is being prepared for publication regarding those results. Further collaborations assessing the genetic susceptibility to PAHs from traffic source and from air pollution (TSP) have been conducted and some associations have been observed. These data and epidemiologic evidence were used for further analyses, based on the models of geospatial lifeline for the estimation of lifetime residential exposures to PAHs and breast cancer risk.

KEY RESEARCH ACCOMPLISHMENTS

- A paper on geographic clustering of residence in early life and subsequent risk of breast cancer was published. (Han, D. Rogerson, PA. Nie, J. Bonner, MR. Vena, JE. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. Geographic Clustering of Residence in Early Life and Subsequent Risk of Breast Cancer. Cancer Causes and Controls 2004 15:921-929)

- A manuscript was published regarding the clustering of lifetime residence and breast cancer risk using exploratory spatial analysis tools based on these lifetime residential history data (Han, D. Rogerson, PA. Bonner, MR. Nie, J.Vena, JE. Muti, P. Trevisan, M. Freudenheim, JL. Assessing Spatio-Temporal Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer. International Journal of Health Geographics 2005 4:9)

- We completed updating of lifetime residential history of breast cancer cases and controls in western New York; All Erie and Niagara county residential location were identified and geocoded, and these were merged into one database. We checked consistency of geocoded addresses in different time points for each individual, updated incomplete addresses using Polk index searches, and validated the consistency of reported years of moved in and out of the residence.

- We completed collection of PAH exposure data from historic traffic information and air pollution sources. A GIS-based traffic model was established to estimate historical residential exposure to PAHs from traffic. An abstract was presented by a colleague.

- A manuscript has been submitted regarding effects of lifetime adult weight change on pre- and post-menopausal breast cancer risk using breast cancer case-control data in western New York. An abstract based on this work was presented at the annual meeting of the Society for Epidemiologic Research in Salt Lake City, June 2004. (American Journal of Epidemiology Supplement 159: S13)

- I have examined clustering of early life residence and breast cancer risk by genetic and biological characteristics, and a poster was presented at the annual meeting of the American Association for Cancer Research in Anaheim , CA, April 2005. (Proc. Amer Assoc Cancer Res 46: 3214)

- I gave an oral presentation for the paper entitled “Clustering of cases by early life residence: evidence for early life environmental exposures in the etiology of breast cancer?” at the Fourth Era of Hope Meeting for the Department of Defense Breast Cancer Research Program in Philadelphia, PA. June 2005.

- I gave an oral presentation for the paper entitled “Assessing the variability of risk surfaces using residential history data in a case control study of breast cancer” at the annual meeting of the International Society for Environmental Epidemiology in New York City, New York, August, 2004. (Epidemiology Supplement 15: S529)

- My paper was selected as one of ten finalists for the Nystrom competition of the Association of American Geographers (AAG), and the paper was presented at the centennial meeting of the AAG, Philadelphia, PA. March, 2004.

- I gave a special guest lecture on “GIS Applications in Environmental Epidemiologic Studies” for a course in the Department of Social and Preventive Medicine, Environmental Epidemiology (SPM551), University at Buffalo-SUNY. April 2005.

- I presented a seminar “Geographical Epidemiology of Breast Cancer in Western New York” as part of the seminar series of the Department of Social and Preventive Medicine, University at Buffalo-SUNY. March, 2004

- I gave an invited talk on “GIS, Medical geography, and Spatial Epidemiology” in Kyung Hee University, Seoul, Korea. June, 2004.

- As a part of my training, I took an epidemiologic method course, “SPM 502: Advanced epidemiologic methods” in the Department of Social and Preventive Medicine, University at Buffalo. Spring, 2004, 3 credit hours, and also took a cancer oncology course, “RPN 530: Oncology for scientists” at the Roswell Park Cancer Institute. Fall, 2004, 3 credit hours.

- I participated in a Harvard School of Public Health workshop “The Public Health Disparities Geocoding Project**”** in June 2005.

- I participated in the NCI Summer Curriculum in Cancer Prevention, and took the course “Principles and Practice of Cancer Prevention and Control” in July, 2003.

- I was selected as one of the participants in Center for Spatially Integrated Social Science (CSISS) summer workshop, “Geographically weighted regression and associated statistics” in August, 2003.

REPORTABLE OUTCOMES

Abstracts/Presentations

- "Assessing the Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer" abstract presented at the Annual Meeting of the International Society for Environmental Epidemiology, New York, NY, August 2004. (abstract attached)

- "Effects of Lifetime Weight Gain on Breast Cancer Risk" abstract presented at the annual meeting of the Society for Epidemiologic Research, Salt Lake City, UT. June 2004. (abstract attached)

- "Geographic Clustering of Residence in Early Life and Risk of Breast Cancer" paper presented at the Annual Meeting of the Association of American Geographers, Philadelphia, PA. March 2004. (abstract attached)

- "Clustering of Cases by Early Life Residence: Evidence for Early Life Environmental Exposures in the Etiology of Breast Cancer?" abstract presented at The fourth Era of Hope meeting for the Department of Defense Breast Cancer Research Program, Philadelphia, PA. June 2005. (abstract attached)

- "Geographic Differences in Breast Cancer Cases and Controls by Genetic and Biological Characteristics: Explaining Clustering of Breast Cancer at Place of Birth" abstract presented at the Annual Meeting of the American Association for Cancer Research. Anaheim, CA. April 2005. (abstract attached)

- "Environmental exposure to traffic polycyclic aromatic hydrocarbons (PAHs) and risk of breast cancer" abstract presented at the Annual Meeting of the American Association for Cancer Research. Anaheim, CA. April 2005. (abstract attached)

- "Traffic Polycyclic Aromatic Hydrocarbons Genetic Susceptibility and Risk of Breast Cancer" abstract presented at the International Congress on Developmental Origins of Health and Disease. Toronto, Canada. (abstract attached)

Publications

- Han, D. Rogerson, PA. Nie, J. Bonner, MR. Vena, J. Vito, D. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. 2004, "Geographic Clustering of Residence in Early Life and Subsequent Risk of Breast Cancer" Cancer Causes and Control 15: 921-929. (manuscript attached)

- Han, D. Rogerson, PA. Bonner, MR. Nie, J.Vena, JE. Muti, P. Trevisan, M. Freudenheim, JL. 2005. Assessing Spatio-Temporal Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer. International Journal of Health Geographics 4:9. (manuscript attached)

- Han, D. Nie, J. Bonner, MR. McCann, SE. Muti, P. Trevisan, M. Ramirez, F. Vito, D. Freudenheim, JL. Lifetime Adult Weight Gain, Central Adiposity, and the Risk of Pre- and Postmenopausal Breast Cancer in the Western New York Exposure and Breast Cancer (WEB) Study. Submitted to International Journal of Cancer. (manuscript attached)

Geographic Differences in Breast Cancer Cases and Controls by Genetic and Biological Characteristics: Explaining Clustering of Breast Cancer at Place of Birth. Han, D. Freudenheim, JL. Nie, J. Bonner, MR. Muti, P. Trevisan, M. Vito, D. Edge, SB. Luyegu, K. Shields, P. Annual Meeting of the American Association for Cancer Research. Anaheim, CA. April 2005.

There is growing evidence that early environmental exposures may be related to risk of breast cancer. We had identified previously geographic clustering of residences in early life, especially at birth, in relation to breast cancer risk. Because women with glutathione-S-transferase M1 (GSTM1) null genotypes may be more sensitive to environmental insults, we report here on clustering by GST genotype. Further, we examined whether there were differences by estrogen receptor (ER) or progesterone receptor (PR) status. We conducted a population-based case control study of incident, primary, histologically-confirmed breast cancer with controls frequency matched to cases on age, race and county of residence (the WEB study). All participants provided lifetime residential histories; we report here on place of birth. ER and PR status were obtained from medical records, and GST genotype was determined by PCR and agarose gel resolution. Geographic differences in clustering of breast cancer cases and controls were examined by the k-function method, a test for general tendency of spatial clustering, in groups stratified by GST genotype and by ER and PR status. We found evidence that breast cancer cases with GSTM1 null genotype were more clustered than controls, while there was no such evidence among GSTM1 wild genotype. However, there was no indication that cases with GSTM1 null genotype were more clustered when compared to cases with the wild genotype. We also observed a tendency to clustering among those with ER+ tumors relative to controls, but not among other groups defined by either ER or PR status. Clustering of women with GSTM1 null genotype may indicate that breast cancer cases with GSTM1 null genotype are more likely to share common environmental exposures at place of birth, but not necessarily congregated in a specific geographic area, and that such genetic susceptibility to early environmental exposure may play a role in subsequent risk of breast cancer. The tendency for clustering of ER positive tumors may also indicate that there are effects of common exogenous exposures related to place of birth among ER positive tumors. These findings are provocative in providing an indication that exogenous exposures at the time of birth may affect breast cancer risk.

Clustering of Cases by Early Life Residence: Evidence for Early Life Environmental Exposures in the Etiology of Breast Cancer? Daikwon Han, Jo L. Freudenheim, Matthew Bonner, Jing Nie, Dominica Vito, Paola Muti, Maurizio Trevisan, Christine Ambrosone, Stephen Edge, Peter Shields. The Fourth Era of Hope Meeting for the Department of Defense Breast Cancer Research Program, Philadelphia, PA. June 2005

The role of environmental exposures that may lead to breast cancer is of continuing interest in breast cancer epidemiology. There is growing evidence that early life exposures may be of significance in the etiology of this disease. While there have been studies of breast cancer clusters based on current address, there have been no such evaluations of clustering of early life residence. Since clustering of breast cancer in space and time may be indicative of potential interactions between exogenous exposures and the subsequent risk of breast cancer, we had previously adapted a novel approach, Geographic Information Systems (GIS) and spatial analysis methods, to identify geographic clustering of residences in early life. Further, we examined whether clustering of residence in early life differed by genetic susceptibility.

We conducted a case control study of incident, primary, histologically confirmed breast cancer with controls frequency matched to cases on age, race and county of residence (the WEB study). All participants provided lifetime residential histories; we identified residential location of participants at the time of their birth, menarche, and their first birth. Glutathione-S-transferase (GST) genotype was determined by PCR and agarose gel resolution. GIS-based spatial clustering analyses were used to identify geographic differences in breast cancer cases and controls in groups stratified by menopausal status and by genotype.

There was a general tendency of geographic clustering for cases for these time periods, and the evidence for clustered residences at birth and at menarche was stronger than that for first birth, especially among premenopausal women. We also found evidence that cases with GSTM1 null genotype were more clustered than controls for these time periods, and that the clustering was stronger at birth and at first birth. However there was no such evidence among those with GSTM1 wild genotype when compared to controls, nor among cases with GSTM1 null genotype when compared to cases with the wild genotype. Analyses by both menopausal status and genotype were not included due to statistical instability.

Clustering of residence in early life, especially among cases with GSTM1 null genotype, may indicate that they are more likely to share common environmental exposures, and that such genetic susceptibility to early environmental exposures may play a role in subsequent risk of breast cancer. By incorporating GIS-based spatial analysis methods, we have shown that there may be effects of common exogenous exposures related to residence in early life, and that such interaction may be important in further assessment of early environmental exposures in relation to breast cancer risk. These findings are provocative in providing an indication that early exogenous exposures may affect the subsequent risk of breast cancer.

Environmental Exposure to Traffic Polycyclic Aromatic Hydrocarbons (PAHs) and Risk of Breast Cancer. Jing Nie, Jan Beyea, Matthew R. Bonner, Daikwon Han, John E. Vena, Peter Rogerson, Dominica Vito, Paola Muti, Maurizio Trevisan, Jo L. Freudenheim. Annual Meeting of the American Association for Cancer Research. Anaheim, CA. April 2005.

Polycyclic aromatic hydrocarbons (PAHs) are an important component of air pollution and potential human carcinogens. While they have been shown to cause mammary cancer in animal studies, the association between PAH exposure and breast cancer risk is not well understood. Traffic emissions are one of the major sources of PAH exposure in cities. Further, growing evidence suggests that there may be critical time periods of exposure in breast cancer initiation and development. In this study, we examined the association between breast cancer risk and exposure to PAHs from traffic emissions estimated for each woman at menarche, at the time when she had her first pregnancy and birth, and at 20 and 10 years prior to interview, using data collected from the Western New York Exposures and Breast Cancer (WEB) study, a population based case control study in western New York. All participants were women, aged 35-79, residents of Erie and Niagara Counties. Cases had incident, primary, histologically-confirmed breast cancer. Controls were randomly selected and frequency-matched to cases on age, race and county. In-person interviews were used to collect data on potential breast cancer risk factors including self-reported lifetime residential history. Traffic volumes on roads were obtained from historical records for the years from 1960-2002. Tailpipe emission data were based on previous reports, including measurements carried out in tunnels or on individual vehicles run in place on test beds. A geographic model, developed by Dr. Beyea and colleagues from the Long Island Breast Cancer project, was used to reconstruct historical traffic PAH, using BaP as a surrogate for total PAH exposure. Cruise emissions, cold engine emissions and intersection emissions were used to estimate total traffic PAH emissions. Meteorological information was also utilized in the geographic dispersion model to assign PAH exposure at each residence. The model was validated using data collected from both Long Island and our study area. We found evidence that higher exposure to traffic PAH emissions at menarche was associated with increased risk of premenopausal breast cancer (OR 2.07, 95% CI 0.91-4.72, p for trend 0.03) and emissions at the time of a woman's first birth was associated with postmenopausal breast cancer (OR 2.58, 95% CI 1.15-5.83, p for trend 0.19). Both associations were limited to lifetime non-smokers. There was no association of traffic emissions with risk for any of the other time periods. These findings provide evidence for both the potential importance of early exposures and the potential importance of an environmental agent in risk of breast cancer.

Traffic Polycyclic Aromatic Hydrocarbons (PAHs) Genetic Susceptibility and Risk of Breast Cancer. Nie, J. Beyea, J. Bonner, MR. Han, D. Vena, J. Rogerson, P. Vito, D. Muti, P. Trevisan, M. Freudenheim, JL. The 3rd International Congress on Developmental Origins of Health and Disease. Toronto, Canada. 2005.

Background: Growing evidence suggests that there may be critical time periods of exposure in breast cancer initiation and development. Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous and exist in the ambient environment at low levels. We previously found evidence that exposure to PAHs based on an estimate of exposure to traffic emissions in a woman's earlier life may be associated with breast cancer risk in adulthood. Glutathione S-transferase mu, a phase II enzyme, is involved in the detoxification of PAHs. There is a common GSTM1 genetic polymorphism that is a deletion of the entire gene. The GSTM1 null genotype is associated with a deficient detoxifying enzyme activity. In this study, we examined the association between GSTM1 genotypes and breast cancer risk, and interaction with traffic emission-PAH exposure estimated for each woman at menarche, at the time when she had her first birth, and at 20 and 10 years prior to interview, using data collected from the Western New York Exposures and Breast Cancer (WEB) study, a population-based case control study. Methods: All participants were women, aged 35-79, residents of Erie and Niagara Counties. Cases had incident, primary, histologically-confirmed breast cancer. Controls were randomly selected and frequency-matched to cases on age, race and county. In-person interviews were used to collect data on potential breast cancer risk factors including self-reported lifetime residential history. Blood samples were collected at the time of the interview and used to determine GSTM1 genotype. A geographic model was used to reconstruct historical traffic PAH exposure at each residence. Results: There was no main effect of GSTM1 on breast cancer risk. While we had previously found an association between higher exposure to traffic emission PAHs and breast cancer risk, we now found evidence that the association was limited to women with GSTM1 null genotype. For exposure at menarche, limited to women living within 250 meters of a road with traffic counts, the upper quartile of PAH exposure was associated with increased risk of premenopausal breast cancer (OR 4.64, 95% CI 0.98-21.94; p for trend 0.01) and emissions at the time of a woman's first birth was associated with increased risk of postmenopausal breast cancer (OR 3.27, 95% CI 0.99-10.84, p for trend 0.02). There was no association of traffic emissions with risk among women with GSTM1 wild-type, or for any of the other time periods. Conclusions: Our findings suggest that there is increased risk of breast cancer associated with exposure to traffic emission PAHs in early life, and that the association is limited to women with GSTM1 null genotype.

Effects of Lifetime Weight Gain on Breast Cancer Risk. D Han, P Muti, M Trevisan, J Nie, D Vito, S Edge, J Freudenheim. Annual Meeting of the Society for Epidemiologic Research, Salt Lake City, UT. 2004.

While there is quite consistent data regarding increased risk of postmenopausal breast cancer with increased body mass index, there is now accumulating data that would indicate that weight gain in adult life is more predictive of breast cancer risk. In this study, we investigated effects of lifetime weight gain on pre- and postmenopausal breast cancer, and effects of weight changes at specific time points in a woman's life. A population-based case control study, the Western New York Exposures and Breast Cancer Study(the WEB study) was conducted. Included were 1,170 women with primary, histologically confirmed, incident breast cancer and 2,116 controls frequency-matched on age and race. Participants were asked to recall their body weight for each decade of their lives from age 20 to the present. Total lifetime weight gain, the difference between weight one year before interview and weight at age 20, and weight changes between each decade were examined. Unconditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI). An increased risk of breast cancer was found for postmenopausal women in the highest (> 27 kg) compared to the lowest quartile (< 9.9 kg) of total lifetime weight gain (adjusted OR 3.28, 95% CI 2.16-4.98). Risk estimates increased in magnitude for increasing time periods of weight gain. Among the premenopausal women, lifetime weight gain was not associated with an increased risk of breast cancer. This study confirms previous findings of increased risk of breast cancer associated with adult weight gain among post- but not premenopausal women, and suggest that weight gain over longer periods of time is associated with higher risk.

Assessing the Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer. Daikwon Han, Peter A Rogerson, Matthew R Bonner, Jing Nie, John E Vena, Jo L Freudenheim. Annual meeting of the International Society for Environmental Epidemiology, New York, NY. 2004

Introduction: Residential location has often been used as a measure of environmental exposure in epidemiologic studies. To examine breast cancer risk associated with residential history, we explored the spatio-temporal patterns of risk surfaces using data on lifetime residential history in a case control study of breast cancer. We applied GIS-based exploratory spatial analyses to obtain risk surfaces, and assessed spatio-temporal variability of the risk surfaces.
Methods: A population-based case control study of breast cancer in western New York (the WEB Study) includes data on the lifetime residential history for breast cancer cases and controls. Participants were asked to provide all locations of earlier residences. Density surfaces of cases and controls were obtained using kernel estimation methods, and the standardized difference in density surfaces was identified to depict elevated areas of breast cancer risk. The significance of the resulting risk surfaces was tested and reported as *p*-values. These surfaces were compared for premenopausal and postmenopausal women. To assess the variability of risk surfaces in space and time, the standardized difference in density surfaces was obtained for specific time periods in a woman's life and for each decade, from 1940s to 1990s.
Results: We found strong evidence of clustering of lifetime residence for premenopausal women (for cases relative to controls), and little evidence of such clustering for postmenopausal women. We also identified the time points contributing most significantly to this result. When density surfaces between cases and controls were compared at each time point, we observed that the earlier decades and early time points, such as residence at birth and menarche, were more likely to be influential time points in understanding overall patterns, relative to the importance of later time points.
Discussion: We were able to pinpoint geographic areas with higher risk, and to assess temporal variability of the risk surfaces by identifying the role of early exposures through exploratory spatial analyses.

Geographic Clustering of Residence in Early Life and Risk of Breast Cancer. Daikwon Han, Annual meeting of the Association of American Geographers, Philadelphia, PA. 2004.

This study examines breast cancer risk associated with lifetime residential history using GIS-based exploratory spatial analyses. Data on residential history and risk factors were collected as part of a population-based case control study of breast cancer in western New York. We were interested in whether there was clustering of breast cancer based on residential location in early life. Under the hypothesis that early exposures may be related to risk of breast cancer, *k*-function differences between breast cancer cases and controls were obtained. We found a general tendency of spatial clustering for residence in early life, compared with the simulated theoretical distribution of expected patterns. The evidence for clustered residential location at birth and at menarche was stronger than that for first birth or other time periods in adult life. Second, relative risk surfaces of cases and controls were obtained to depict elevated areas of breast cancer risk using kernel smoothing methods. We observed stronger evidence of geographic clustering of lifetime residence for pre-menopausal women relative to that for post-menopausal women, and clustering of early-life residence relative to that of adult-life residence. Our findings suggest that there may be identifiable etiological processes linking exposure and breast cancer risk, especially for pre-menopausal women, and that early exposures may be of particular importance. This study provides additional evidence that early environmental exposures may be related to breast cancer risk.

CONCLUSIONS

We have now completed collection of historic traffic information, and completed updating of the lifetime residential histories for our dataset for the breast cancer cases and controls. The development of a theoretical framework measuring similarity and difference of individual's lifetime residential history was also completed, and a GIS-based traffic model was established. Because we found evidence of association between PAHs exposures in early life in relation to breast cancer risk, these data and epidemiologic evidence were used for further analyses, based on the models of geospatial lifeline for the estimation of lifetime residential exposures to PAHs and breast cancer risk. Findings from these studies indicate that environmental exposures in early life may be associated with breast cancer risk. Since we found evidence of association between PAHs exposures in relation to breast cancer risk, especially PAH exposures during sensitive time periods in early life, we strongly believe that models of geospatial lifeline can be effectively used for the estimation of lifetime residential exposures to PAHs and breast cancer risk.

REFERENCES

1. Ahlgren M, Melbye M, Wohlfahrt J, Sorensen T. Growth patterns and the risk of breast cancer in women. N Engl J Med 2004;351:1619-26.
2. Bailey TC, Gatrell AC. Interactive Spatial Data Analysis. Harlow, England: Longman Scientific & Technical; 1995.
3. Ballard-Barbash R. Anthropometry and breast cancer. body size-a moving target. Cancer 1994;74:1090-100.
4. Barker, DJP. Fetal and infant origins of adult disease. London: BMJ Publishing; 1992.
5. Berkey C. Frazier A. Gardner J. Colditz G. Adolescence and breast carcinoma risk. Cancer 1999;85:2400-9.
6. Bonner, MR. Han, D. Nie, J. Rogerson, PA. Vena, JE. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. Breast Caner Risk and Exposure in Early Life to Polycyclic Aromatic Hydrocarbons Using Total Suspended Particulates as a Proxy Measure. Cancer Epidemiology, Biomarkers and Prevention 2005 14: 53-60.
7. Bonner, MR. Nie, J. Han, D. Rogerson, PA. Vena, JE. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. Environmental Tobacco Smoke Exposure in Early Life and the Risk of Breast Cancer. Cancer Causes and Control. 2005 16:683-89.
8. Bowman AW, Azzalini A. Applied Smoothing Techniques for Data Analysis. Oxford: Clarendon Press; 1997.
9. Brody JG, Rudel RA. Environmental pollutants and breast cancer. Environ Health Perspect 2003 111: 1007-1019.

10. Colditz, GA. Frazier, AL. Models of breast cancer show that risk is set by events of early life: prevention efforts must shift focus. Cancer Epidemiology, Biomarkers, and Prevention 1995; 4:567-571.
11. Cressie, N. Statistics for spatial data. New York: Wiley; 1993.
12. Cromley EK, McLafferty SL. GIS and Public Health. New York: The Guilford Press; 2002.
13. DeBruin LS, Josephy PD. Perspectives on the chemical etiology of breast cancer. Environ Health Perspect 2002 110(suppl): 119-128.
14. Diggle PJ. Statistical Analysis of Spatial Point Patterns. London: Academic Press; 1983.
15. Elliott P, Wakefield JC, Best NG, Briggs, DJ, eds. Spatial Epidemiology: Methods and Applications. Oxford: Oxford University Press; 2000.
16. Friedenreich C. Review of anthropometric factors and breast cancer risk. Eur J Cancer Prev 2001;10:15-32.
17. Han, D. Rogerson, PA. Bonner, MR. Nie, J.Vena, JE. Muti, P. Trevisan, M. Freudenheim, JL. Assessing Spatio-Temporal Variability of Risk Surfaces using Residential History Data in a Case Control Study of Breast Cancer. International Journal of Health Geographics 2005 4:9.
18. Han, D. Rogerson, PA. Nie, J. Bonner, MR. Vena, JE. Muti, P. Trevisan, M. Edge, S. Freudenheim, JL. Geographic Clustering of Residence in Early Life and Subsequent Risk of Breast Cancer. Cancer Causes and Control 2004 15: 921-929.
19. Harvie M, Howell A, Vierkant RA, Kumar N, Cerhan JR, Kelemen LE, et al. Association of gain and loss of weight before and after menopause with risk of

postmenopausal breast cancer in the Iowa Women's Health study. Cancer Epidemiol Biomarkers Prev 2005 14:656-61.

20. Huang Z. Hankinson SE. Colditz GA. Stampfer MJ. Hunter DJ. Manson JE, et al. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997;278:1407-11.
21. Hunter DJ, Willett WC. Diet, body size, and breast cancer. Epidemiol Rev 1993;15:110-32.
22. International Agency for Research on Cancer (IARC). Handbooks of cancer prevention: weight control and physical activity. Lyon: IARC Press;2002.
23. Jacquez GM, Goovaerts P, Rogerson PA. Space-time intelligence systems: technology, applications and methods. Journal of Geographical Systems 2005 7:1-5.
24. Kawachi I, Berkman LF (eds). Neighborhoods and Health. New York: Oxford University Press; 2003.
25. Khan OA, ed. Geographic Information Systems and Health Applications. Hershey, PA: Idea Group Publishing; 2003.
26. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: The Public Health Disparities Geocoding Project. Am J Epidemiol 2002; 156:471-482.

27. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. Am J Public Health 1992; 82:703-710.
28. Kuh, D. Ben-Shlomo, Y. A life course approach to chronic disease epidemiology: tracing the origins of ill-health from early to later life. Oxford: Oxford University Press; 1997.
29. Kumar NB. Lyman GH. Allen K. Cox CE. Schapira DV. Timing of weight gain and breast cancer risk. Cancer 1995;76:243-49.
30. Laden F. and Hunter DJ. Environmental risk factors and female breast cancer. Annual Review of Public Health 1998; 19:101-123.
31. Lawson A, Biggeri A, Bohning D, Lesaffre E, Viel J, Bertollini R, eds. Disease Mapping and Risk Assessment for Public Health. Chichester: John Wiley and Sons; 1999.
32. Lewis-Michl EL, Melius JM, Kallenbach LR, Ju CL, Talbot TO, et al. Breast cancer risk and residence near industry or traffic in Nassau and Suffolk Counties, Long Island, New York. Arch Environ Health 1996 51: 255-65.
33. Mark, D.M. Egenhofer, M.J. Bian, L. Hornsby, K.E. Rogerson, P.A. Vena, J. Spatiotemporal GIS analysis for environmental health: solutions using geospatial lifelines. GEOMED'99: Proceedings of the International Workshop on Geomedical Systems; 2000.
34. New York State Department of Health. The New York State Cancer Surveillance Improvement Initiative. Albany, NY; 1998.

35. Paulu C, Aschengrau A, Ozonoff D. Exploring associations between residential location and breast cancer incidence in a case-control study. Environ Health Perspect 2002 110: 471-478.
36. Potischman N, Swanson CA, Siiteri P, Hoover RN. Reversal of relation between body mass and endogenous estrogen concentrations with menopausal status. J Natl Cancer Inst 1996;88:756-58.
37. Potischman, N. and Troisi, R. In-utero and early life exposures in relation to risk of breast cancer. Cancer Causes and Control 1999; 10:561-573.
38. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. Geographic information systems and public health: mapping the future. Public Health Rep 1999; 114:359-373.
39. Ripley, BD. Spatial statistics. New York: Wiley; 1980.
40. Rothman KJ. A sobering start for the cluster buster's conference. Am J Epidemiol 1990 132: S6-13.
41. Rushton G. Improving the geographic basis of health surveillance using GIS. In: GIS and Health (Gatrell AC, Löytönen M, eds). London: Taylor and Francis, 1998, 63-80.
42. Sabel CE, Gatrell AC, Loytonen M, Maasilta P, Jokelainen M. Modelling exposure opportunities: estimating relative risk for motor neurone disease in Finland. Soc Sci Med 2000 50: 1121-37.
43. Silverman BW. Density Estimation for Statistics and Data Analysis. London: Chapman and Hall; 1986.

44. Sinha, G. Mark, D. Measuring similarity between geospatial lifelines in studies of environmental health. Journal of Geographical Systems 2005 7:115-136.

45. Stoll BA. Adiposity as a risk determinant for postmenopausal breast cancer. Int J Obes Relat Metab Disord. 2000;24:527-33.

46. Stoll BA. Timing of weight gain in relation to breast cancer risk. Ann Oncol 1995;6:245-8.

47. Trichopoulos, D. Hypothesis: does breast cancer originate in-utero? Lancet 1990; 335:939-940.

48. Venables WN, Ripley BD. Modern Applied Statistics with S-Plus. New York : Springer; 1998.

49. Ward MH, Nuckols JR, Weigel SJ, Maxwell SK, Cantor KP, Miller RS. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System. Environ Health Perspect. 2000 108: 5-12.

50. Wartenberg D, Greenberg M, Lathrop R. Identification and characterization of populations living near high-voltage transmission lines: a pilot study. Environ Health Perspect 1993 101: 626-632.

51. Wartenberg D. Investigating disease clusters: why, when and how ? J R Statist Soc A 2001 164:13-22

52. Willett WC. Nutritional Epidemiology. Oxford:Oxford University Press;1998.

53. Wolff, MS. Collman, GW. Barrett, C. and Huff, J. Breast cancer and environmental risk factors: epidemiologic and experimental findings. Annual Review of Pharmacology and Toxicology 1996; 36:573-596.

APPENDIX

Cancer Causes and Control **15**: 921–929, 2004.

Geographic clustering of residence in early life and subsequent risk of breast cancer (United States)

Daikwon Han[1,2,*], Peter A. Rogerson[2,3], Jing Nie[1], Matthew R. Bonner[1], John E. Vena[4], Dominica Vito[1], Paola Muti[1], Maurizio Trevisan[1], Stephen B. Edge[5] & Jo L. Freudenheim[1]
[1]*Department of Social and Preventive Medicine, University at Buffalo, Buffalo, NY 14214, USA;* [2]*Department of Geography and National Center for Geographic Information and Analysis, University at Buffalo, Buffalo, NY 14261, USA;* [3]*Department of Biostatistics, University at Buffalo, Buffalo, NY 14214, USA;* [4]*Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC 29208, USA;* [5]*Department of Breast and Soft Tissue Surgery, Roswell Park Cancer Institute, Buffalo, NY 14263, USA*

Received 23 March 2004; accepted in revised form 3 June 2004

Key words: breast cancer, early-life exposure, spatial clustering.

Abstract

Objective: This study focused on geographic clustering of breast cancer based on residence in early life and identified spatio-temporal clustering of cases and controls.
Methods: Data were drawn from the WEB study (Western New York Exposures and Breast Cancer Study), a population-based case–control study of incident, pathologically confirmed breast cancer (1996–2001) in Erie and Niagara counties. Controls were frequency-matched to cases on age, race, and county of residence. All cases and controls used in the study provided lifetime residential histories. The *k*-function difference between cases and controls was used to identify spatial clustering patterns of residence in early life.
Results: We found that the evidence for clustered residences at birth and at menarche was stronger than that for first birth or other time periods in adult life. Residences for pre-menopausal cases were more clustered than for controls at the time of birth and menarche. We also identified the size and geographic location of birth and menarche clusters in the study area, and found increased breast cancer risk for pre-menopausal women whose residence was within the cluster compared to those living elsewhere at the time of birth.
Conclusion: This study provides evidence that early environmental exposures may be related to breast cancer risk, especially for pre-menopausal women.

Introduction

Breast cancer is one of the leading causes of death among women in the United States. However, the epidemiology of breast cancer is not yet fully understood. We also do not fully understand mechanisms for the known risk factors; for instance, why changes in age at menarche or age at first birth have an impact on breast cancer risk. A substantial degree of geographical variation in breast cancer incidence and mortality in the US has been observed [1, 2]. While inconclusive, several environmental risk factors are also believed to be involved in breast cancer incidence [3, 4]. There is speculation that environmental factors may explain geographic variation in breast cancer rates not explained by known risk factors. For this reason, the potential role of environmental exposures in breast cancer risk is of particular interest.

In addition, there is a growing interest in early life and lifetime exposures in relation to breast cancer risk. The life course approach is of interest because there may be sensitive time periods for exposures and/or there may be cumulative effects of lifetime exposure involved in breast cancer incidence [5, 6]. Early life has an effect on breast cancer etiology evidenced by the known risk factors

*Address correspondence to: Daikwon Han, Department of Social and Preventive Medicine, University at Buffalo, 3435 Main St., Farber Hall, Rm 270, Buffalo, NY 14214, USA. Ph.: +1-716-829-2975 ext. 605; Fax: +1-716-829-2979; e-mail: dhan@buffalo.edu

such as age at menarche, age at first birth and parity. There is new evidence that even earlier exposures may have an impact on adult breast cancer risk [7]. Trichopoulos [8] suggested that the *in-utero* and perinatal period might be pathologically significant and that the risk of adult breast cancer could be related to high estrogen exposure in early life. There is also accumulating evidence that factors related to early exposure, such as birthweight, may be related to risk [9, 10].

There has been little research investigating possible effects of environmental exposures in early life on subsequent breast cancer risk. Using residence as a proxy measure for environmental exposures, we investigated whether there was any evidence of geographic clustering of adult breast cancer cases associated with their residences in early life. Clustering analyses have often been used to provide clues for the unknown etiology of disease, and thus to generate hypotheses for further epidemiologic research [11]. We looked at the geographic clustering of residence at early critical time points: at birth, at menarche, and at the woman's first birth. By comparing differences in clustering patterns between case and control residences, we were interested in identifying time periods critical to potential environmental exposures and subsequent breast cancer risk.

Methods

Population-based case–control study of breast cancer

We conducted a case–control study of breast cancer in western New York – the WEB study (Western New York Exposures and Breast Cancer Study) . Cases were women, age 35–79 with incident, primary, pathologically confirmed breast cancer diagnosed in Erie and Niagara counties during the period 1996–2001, with no previous cancer diagnosis other than non-melanoma skin cancer. Controls were frequency matched to cases on age, race, and county of current residence; controls under 65 years of age were randomly selected from a New York State Department of Motor Vehicles list and those 65 years and over were chosen from a Health Care Finance Administration list. We ascertained cases by having a nurse–case finder visit the pathology departments of almost all hospitals in these counties. One hospital which did not participate does almost no cancer surgery and refers patients to other participating hospitals. For the one other hospital that did not participate, breast cancer cases were identified in the practice of the breast surgeons who see more than 99% of the cases from that hospital. Extensive in-person interviews and self-administered questionnaires were used to ascertain lifetime residential history and other breast cancer risk factors. A total of 1166 cases and 2105 controls were interviewed. Response rates were 72 and 65% for cases and controls, respectively.

All participants were asked to complete a lifetime residential history, to list the street address, town/city and zip code for their current address and then all other previous addresses throughout their lifetime. Participants provided 20,240 addresses, an average of approximately six addresses for each individual. In this study we focused on residence at the time of the participants' birth, menarche, and at the time that she had her first birth. Analyses were restricted to women residing in Erie or Niagara counties at each of these time points. There were, of course, participants whose addresses were the same for two or more of these times.

For women with incomplete residential information, additional information was obtained using historical city directories. We used these directories to find old addresses, and utilized various resources, such as web searches and commercial address databases for recent addresses. We also examined validity and reliability of reports of earlier residences in a number of ways. For birth addresses, we asked for information on birth address twice and have collected information on reliability of response. For the other time periods, we used information on maiden name and partial address information provided by the participants to search for records in city directories for the appropriate time periods. To improve our ability to geocode addresses, we developed several strategies. First, all addresses were standardized to be matched with the standard format used in GIS. We used the enhanced version of TIGER (Topologically Integrated Geographic Encoding and Referencing Systems), GDT/Dynamap 2000 [12], and overall matching rates were improved about 15–20 % when compared with the use of TIGER as a reference theme. We also used the stand-alone address cleaner ZP4 (Semaphore Co.) to correct and update zip code information to be matched with United States Postal Services certified addresses.

More than 85% of addresses were geocoded using the above strategies and resources. We failed to geocode some addresses primarily because of missing residential information, such as missing street numbers or street names. Since we are dealing with historical residential information, the likelihood of missing previous residential information was higher than that for current residential information. Table 1 is a summary table showing the numbers of cases and controls with complete residential information who resided in the two counties for each of the time periods. The percentage of missing residential information associated with

Table 1. Residential history of breast cancer cases and controls: numbers and percentage of complete and missing residences in Erie and Niagara counties: WEB Study, 1996–2001

	Complete residence		Incomplete or missing residence		Total eligible Erie and Niagara county residence at each time period	
	Case	Control	Case	Control	Case	Control
Birth	505 (79.9%)	804 (81.0%)	127 (20.1%)	189 (19.0%)	632	993
Menarche	673 (87.3%)	1143 (88.1%)	98 (12.7%)	154 (11.9%)	771	1297
First birth	616 (86.4%)	1153 (87.3%)	97 (13.6%)	167 (12.7%)	713	1320

each early life event was highest for birth addresses, at about 20%.

Clustering analyses of residences

To compare clustering patterns of breast cancer cases and controls at each time period, the primary method used was based on the k-function [13]. The k-function for a point process is defined as the number of events within distance h of an arbitrary event, divided by the overall intensity of events. It is estimated by

$$\lambda\hat{k}(h) = \sum_{i=1}^{n}\sum_{j=1}^{n} w(s_i, s_j)^{-1} I(d_{ij} \leq h)/n, \quad h > 0$$

where n is the number of events, λ is the expected density of events in the study region, h is the pre-specified distance, d_{ij} is the Euclidian distance between point i and point j, I is an indicator function that is equal to one if inter-event distances (d_{ij}) are less than or equal to h, and zero otherwise, and $w(s_i, s_j)$ is an edge correction estimator which is the proportion of the circumference of a circle centered at s_i, passing through s_j and that is inside the study area A [14]. Under the null hypothesis of spatial randomness, the expected value of $k(h)$ is πh^2. Geographic clustering will yield values of the k-function that are greater than this, since clustering will result in more pairs of points separated by a distance of h than would be expected in a random pattern.

We used the difference between k-functions for cases and controls to compare two patterns (*i.e.,* $D(h) = k_{\text{case}}(h) - k_{\text{control}}(h)$). Positive values of $D(h)$ indicate spatial clustering of cases relative to the spatial clustering of controls. Under the null hypothesis of random labeling of cases and controls, the expected value of $D(h)$ is zero, indicating that the k-functions of the cases and controls are the same. The test statistic, $D(h)$, was calculated with confidence envelopes using the *splancs* library in *S-plus* [15]. We obtained the approximate 95% confidence limits for two standard errors ($\pm 2\sqrt{Var\{D(h)\}}$) at the $\alpha = .05$ level [16]. When the estimated function $D(h)$ deviated from zero by greater than two standard deviations, we interpreted this as a statistically significant difference between the case and control patterns.

We also employed a spatial clustering method to identify significant geographic clusters of breast cancer cases. The spatial scan statistic [17], which considers the likelihood of observing the actual number of cases inside of a circle under the null hypothesis of no clustering, was applied to residence at early life events. We were mainly interested in spatial clustering of high rates, and employed the Bernoulli model based on the locations of individual cases and controls [18]. In addition, odds ratios (OR) and 95% confidence intervals (95% CI) were obtained using logistic regression, adjusting for age, education, age at menarche, parity, history of benign breast disease, family history of breast cancer. All analyses were conducted for the entire group of study participants and for data stratified on menopausal status. Women were considered post-menopausal if their menses had ceased permanently and naturally. Among other women, participants were also considered post-menopausal if any of the following conditions were true: they were on hormone replacement therapy and were over age 55, they had had a bilateral oophorectomy, they had had a hysterectomy without removal of the ovaries and they were older than 50, their menses had ceased permanently due to radiation or other medical treatment and they were older than 55.

Results

Characteristics of subjects included in the analysis, subjects with missing residential information, and subjects excluded due to residence outside of Erie and Niagara counties, are shown in Table 2. About half of the sample was excluded for each time period; the highest percentage of ineligible cases and controls was at the birth residence (46 and 51% respectively). However, we found little difference in characteristics between

Table 2. Characteristics of subjects included in the analysis, subjects with missing residential information, and subjects excluded due to residence outside of the study area (Mean ± SD): WEB Study, 1996–2001

	Cases (n = 1166)			Controls (n = 2105)		
	Included	Missing	Ineligible*	Included	Missing	Ineligible*
Birth	(n = 505)	(n = 127)	(n = 534)	(n = 804)	(n = 189)	(n = 1112)
Age (years)	56.5 ± 10.9	60.0 ± 11.0	58.9 ± 11.3	55.6 ± 11.7	58.0 ± 11.8	59.4 ± 11.7
Education (years)	13.5 ± 2.4	13.1 ± 2.5	13.6 ± 2.7	13.4 ± 2.2	13.2 ± 2.2	13.3 ± 2.5
Parity	2.2 ± 1.5	2.4 ± 1.7	2.4 ± 1.8	2.6 ± 1.8	2.7 ± 1.8	2.8 ± 1.8
Age at menarche (years)	12.4 ± 1.5	12.6 ± 1.5	12.7 ± 1.7	12.7 ± 1.7	12.6 ± 1.6	12.7 ± 1.7
Age at first birth (years)	24.3 ± 4.6	23.5 ± 4.5	24.2 ± 5.1	24.5 ± 4.3	23.5 ± 4.2	24.0 ± 4.7
Pre-menopausal (%)	35.2	18.9	26.4	31.7	28.6	24.6
Body Mass Index	28.2 ± 6.4	28.4 ± 5.8	28.7 ± 6.4	28.0 ± 6.2	28.2 ± 6.0	28.4 ± 6.4
Family history of breast cancer (% yes)	21.3	18.9	20.2	12.7	16.2	12.4
History of benign breast disease (% yes)	34.9	37.0	32.8	22.3	25.9	20.6
Menarche	(n = 673)	(n = 98)	(n = 395)	(n = 1143)	(n = 154)	(n = 808)
Age (years)	56.6 ± 10.7	60.1 ± 11.6	59.5 ± 11.3	56.0 ± 11.7	60.2 ± 11.7	59.9 ± 11.6
Education (years)	13.5 ± 2.4	12.8 ± 2.6	13.6 ± 2.8	13.4 ± 2.2	13.0 ± 2.3	13.3 ± 2.6
Parity	2.2 ± 1.6	2.8 ± 1.8	2.5 ± 1.8	2.6 ± 1.8	2.9 ± 2.1	2.9 ± 1.8
Age at menarche (years)	12.5 ± 1.6	12.8 ± 1.5	12.7 ± 1.7	12.7 ± 1.6	12.6 ± 1.7	12.7 ± 1.7
Age at first birth (years)	24.3 ± 4.6	23.0 ± 4.3	24.2 ± 5.3	24.4 ± 4.5	23.8 ± 4.4	24.0 ± 4.6
Pre-menopausal (%)	30.3	24.5	24.6	33.8	23.4	23.3
Body Mass Index	28.1 ± 6.2	29.5 ± 6.4	28.7 ± 6.5	28.3 ± 6.5	27.6 ± 5.5	28.2 ± 6.1
Family history of breast cancer (% yes)	20.2	22.4	20.6	13.1	13.2	12.1
History of benign breast disease (% yes)	34.5	40.8	31.9	22.3	19.5	21.3
First Birth	(n = 616)	(n = 97)	(n = 453)	(n = 1153)	(n = 167)	(n = 785)
Age (years)	57.4 ± 11.1	58.9 ± 10.8	58.5 ± 11.2	57.0 ± 11.7	60.6 ± 10.7	58.5 ± 12.0
Education (years)	13.4 ± 2.3	13.0 ± 2.9	13.7 ± 2.8	13.3 ± 2.2	13.0 ± 2.1	13.4 ± 2.6
Parity	2.7 ± 1.3	3.1 ± 1.5	1.7 ± 1.9	3.0 ± 1.5	3.4 ± 1.7	2.2 ± 2.0
Age at menarche (years)	12.6 ± 1.5	12.5 ± 1.8	12.6 ± 1.6	12.7 ± 1.6	12.6 ± 1.5	12.7 ± 1.7
Age at first birth (years)	24.8 ± 4.8	22.2 ± 4.1	23.4 ± 4.9	24.7 ± 4.6	22.9 ± 3.5	23.3 ± 4.4
Pre-menopausal (%)	29.4	26.8	26.0	32.2	18.0	26.6
Body Mass Index	28.4 ± 6.3	30.1 ± 6.6	28.2 ± 6.3	28.1 ± 6.1	28.3 ± 6.5	28.4 ± 6.4
Family history of breast cancer (% yes)	21.2	23.7	18.6	11.5	19.4	13.6
History of benign breast disease (% yes)	34.7	37.1	32.7	21.0	25.1	22.0

* Ineligible due to residence outside of Erie and Niagara county.

those subjects included and those subjects with addresses outside of these two counties.

Mapping was used to identify geographic patterns of breast cancer cases and controls for each of the early life events. Maps showing the locations of cases and controls in Figure 1 portray the underlying geographic patterns of breast cancer cases and controls in the study area. The rectangular region was used instead of the actual county boundary as an approximate boundary of the study area to protect individuals' confidentiality. The purpose of such mapping is to inspect patterns visually – the first step in any spatial analysis. Geographic patterns do not appear to vary much from one time period to the next, and they appear to reflect patterns of population distribution in the study area. However, it is difficult to determine whether they were clustered or dispersed relative to population from visual inspection alone, because of the large number of data points.

To assess potential effects of geographic selection bias in our study, we also examined the distribution of current residence in relation to other population data on the geographic distribution of breast cancer cases and the general population. We did not find differences in the geographic distribution of participating and non-participating cases, or between controls and the underlying population, except some tendency for both cases and controls living closer to the interview site to be somewhat more likely to participate than those living further away.

Spatial clustering of residences associated with early life events

We obtained differences between the case and control patterns for locations associated with each early life event. The k-function differences for values of h up to 15 miles, with approximate 95% confidence envelopes, are shown

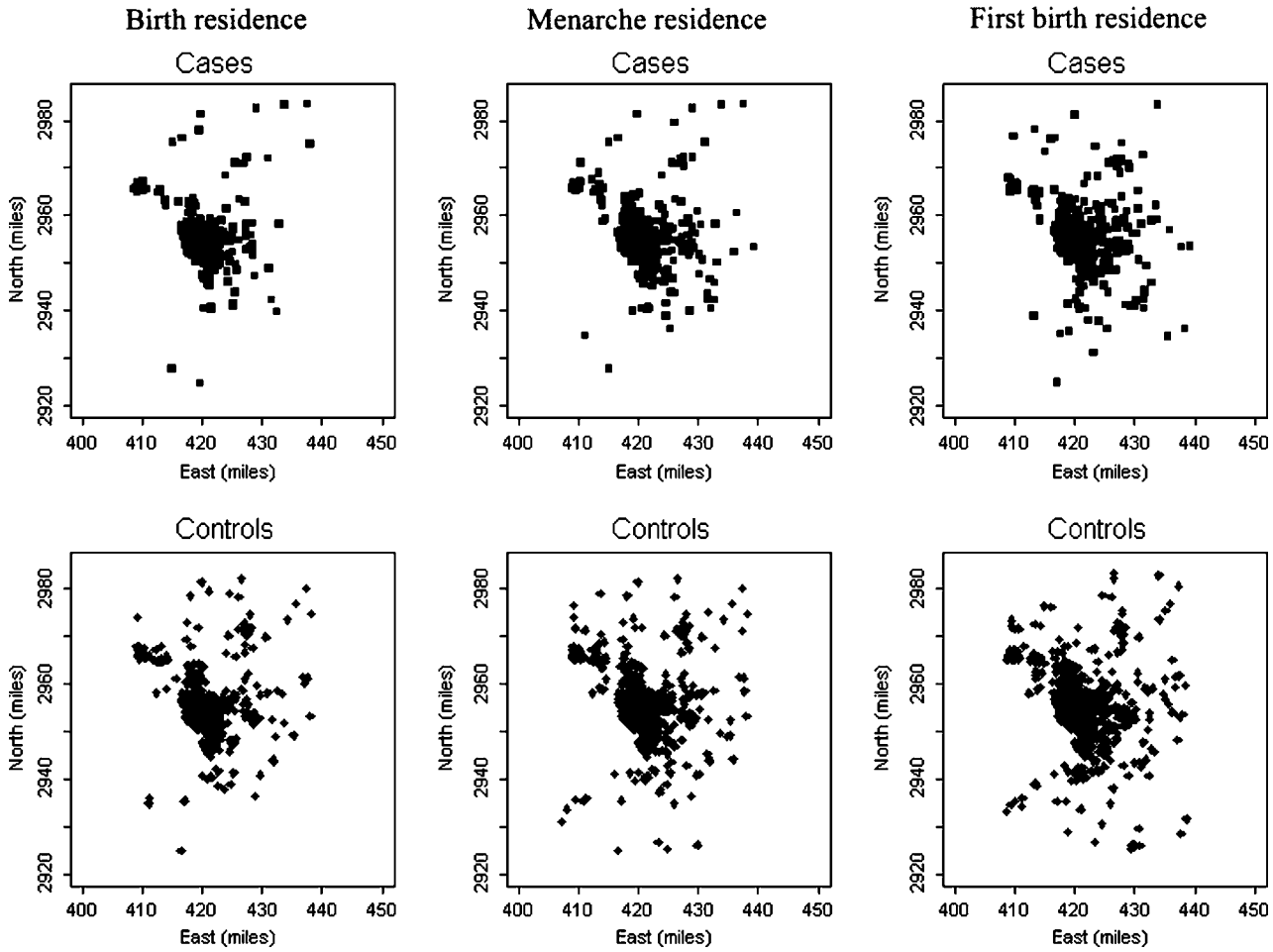

Fig. 1. Residential location of breast cancer cases and controls at each time period: WEB Study, 1996–2001.

in Figure 2. The maximum value of h is generally taken as one-third of the linear extent of the study area [19]. Any patterns beyond this scale can be disregarded, since either peaks or troughs in this geographic scale are difficult to interpret, and are potentially misleading. Figure 2a shows k-function differences for birth residence. It is clear that the estimated function shows strong evidence of spatial clustering, that is, of clustering of cases relative to controls. There was no significant difference up to three miles; statistically significant differences were detected beyond the scale of three miles. There is also evidence of some degree of clustering for breast cancer cases at menarche residence (Figure 2b). Estimates of the D-function are positive but not statistically significant up to seven miles; spatial clustering of breast cancer cases occurs at a scale of about 7–15 miles. For residence at women's first birth and for current residence, the difference is not statistically significant; the plot falls within the confidence interval over all distances (Figures 2c and d).

To determine whether there are any differences in clustering patterns by menopausal status, the k-function difference was performed for pre-menopausal and post-menopausal women separately (Figure 3). We found significant clustering of pre-menopausal breast cancer cases compared to controls for both birth and menarche residence (Figures 3a), while there is no evidence of clustering for post-menopausal breast cancer cases for either period (Figures 3b). We did not find evidence of clustering for first birth and current residence (at diagnosis) for either group (not shown). Estimated functions at birth residence show a strong clustering of pre-menopausal cases over the entire geographic scale with a peak at seven miles. Values are positive for post-menopausal cases, but not statistically significant. For menarche residence, we also observed a strong clustering of pre-menopausal cases with a peak at about 8–10 miles. Again differences are not statistically significant for post-menopausal women at menarche residence.

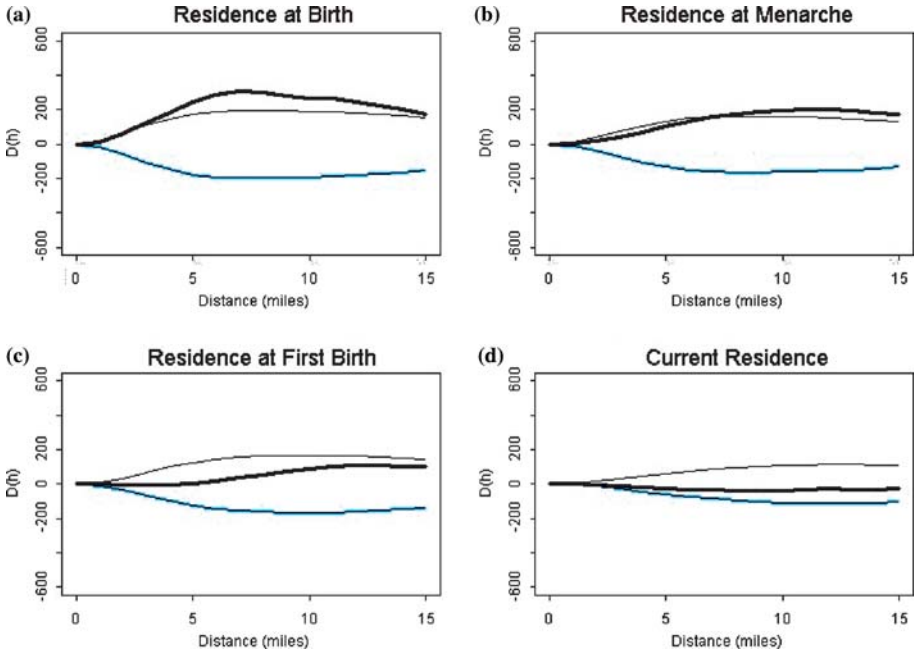

Fig. 2. *k*-function differences in clustering patterns between breast cancer cases and controls, WEB Study, 1996–2001: shown are *k*-function differences in black and 95% confidence limits in grey.

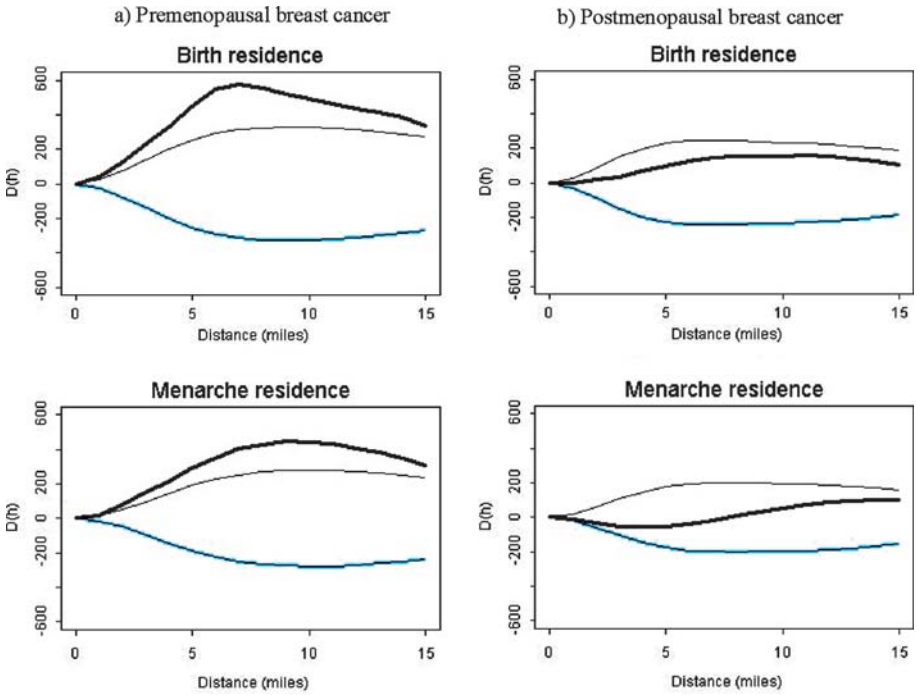

Fig. 3. *k*–function differences in clustering patterns between breast cancer cases and controls by menopausal status, WEB Study, 1996–2001.

Identifying the geographic location of breast cancer clusters

To identify the geographic location of areas with higher intensities for pre-menopausal cases in the study area, the spatial scan statistic was applied to residences of pre-menopausal women at the time of birth and menarche. Maps in Figure 4 present results of the clustering analysis. The circle in Figure 4a indicates

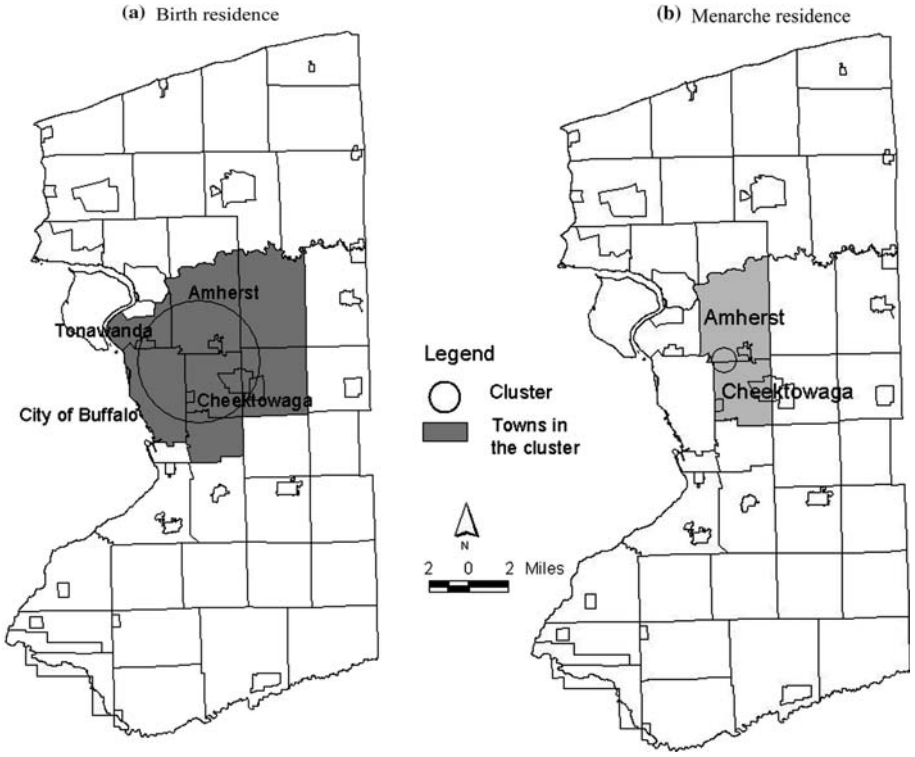

Fig. 4. Geographic clustering of residence at birth and menarche: pre-menopausal breast cancer, WEB Study, 1996–2001.

clustering of birth residence for pre-menopausal cases when compared to controls. We found a circular cluster of birth residence for breast cancer cases with a 5.7-mile radius in the area including part of the city of Buffalo, and the towns of Amherst, Cheektowaga, and Tonawanda (shaded areas). There are 100 observed breast cancer cases inside the cluster, while 76 breast cancer cases are expected. The cluster was significant at <0.01 with 999 Monte-Carlo simulations.

Further, we examined breast cancer risk associated with residence in the cluster at the time of birth. When we compared other breast cancer risk factors, such as age, education, and age at menarche, for the pre-menopausal breast cancer cases and controls whose birth residence was inside the cluster to those who lived outside of cluster, we did not find significant differences between the two groups (data not shown). We observed an elevated breast cancer risk for pre-menopausal women living in the cluster at the time of birth. With subjects living outside the cluster as a reference group, the adjusted odds ratio was 2.65 (95% CI 1.75–4.0) after controlling for age, education, age at menarche, parity, history of benign breast disease, and family history of breast cancer.

We also identified clustering of menarche residence for pre-menopausal women and obtained similar results as for birth residence. We were able to identify a small clustering of menarche residences for pre-menopausal breast cancer cases. A small cluster in the center of those four towns was detected (Figure 4b). It is a cluster with 0.8 mile radius and is statistically significant at $p < 0.05$. The cluster contains nine observed and 3.1 expected breast cancer cases, yielding a relative risk (ratio of observed to expected breast cancer cases) of 2.9. A secondary cluster was also detected near the city of Buffalo. It has a three-mile radius and relative risk of 1.38 with 65 observed and 47 expected breast cancer cases, but it is not statistically significant ($p = 0.38$).

Discussion

To our knowledge, no other studies have examined clustering of residential locations associated with cancer during early life: studies have examined clustering of residential locations at the time of diagnosis or death [20]. Critical time periods, including birth, menarche, and women's first pregnancy, as important early life and reproductive events in women's life, may play a substantial role in the risk of breast cancer. Under the hypothesis that there may be sensitive time periods in women's lives that will carry greater risk for exposure,

the essential question was whether cases were more clustered than the underlying population, as represented by the controls. We found that cases were more clustered than controls at the time of birth and menarche, and it was due to clustering of residence for pre-menopausal, but not for post-menopausal breast cancer. The evidence for clustering of residential locations at birth and menarche was stronger than evidence for clustering at the time of women's first birth or other time periods in adult life. Our findings suggest that there may be identifiable etiological processes linking exposure and breast cancer risk, especially for pre-menopausal women, and that early exposures may be of particular importance.

This study provided a unique opportunity to examine clustering of breast cancer cases and controls at various points during early life. The facts that the study area had a relatively stable population and about 40% of study participants were lifetime residents, made the results more reliable. The evidence that residence in early life was important in the geographical clustering of breast cancer cases may be of particular importance for understanding environmental determinants of breast cancer. These findings suggest the importance of early or lifetime exposure in relation to disease risk in adult life, and also the potential role of the effects of migration on exposures and disease risk. Although migration can have a serious effect on the detection of geographical differences in disease risk, it has not been adequately addressed in previous clustering analyses [21]. Further investigations are required to prove any relationship between geographic clustering of residence and breast cancer risk, and the effects of residential changes on exposures should be considered in these studies.

Our finding of clustering was restricted to pre-menopausal breast cancer. We stratified on menopausal status because of evidence that there were differences in risk factors for pre- and post-menopausal women [22]. The mechanism of the observed difference is not clear. It could be that early life exposures impact pre-menopausal more than post-menopausal disease because of greater temporal proximity. There is some evidence, though not consistent, that other early exposures may differ by menopausal status. For example, there are data suggesting that birthweight may be more associated with pre- than with post-menopausal breast cancer [9, 23].

The results should be interpreted cautiously due to the fact that there may be some artifacts of the analysis. First, it is important to note that spatial point patterns are complex to summarize in a single way [24]. For example, the use of cumulative scales in the application of the *k*-function method may influence the outcome [25]. In particular, clustering is more likely to be detected on a larger geographic scale, and it tends to show continuous patterns over several neighboring scales due to the fact that the geographical scales are cumulative. Further refinement of methods to summarize spatial point patterns may provide more reliable results, as well as more accurate estimates of disease risk.

Second, this study is limited to current residents in the study area because we focused on the residential environment of Erie and Niagara counties; participants residing outside of these two counties at the time of each early life event were not included. The existence of missing residential information and potential selection bias due to non-participation may influence the results. As noted, we found no difference in participation by residence for cases compared to controls. Further we would expect that our findings on the clustering of early-life residence would be less subject to potential geographic selection bias than would current residence. We found a greater degree of clustering for residence at early life than for current residential location.

Further, the fact that residence at birth and menarche were often the same made it difficult to differentiate associations for the two time periods. For 22% of cases and 35% of controls, the menarche residence was the same as their birth residence. While the observed tendencies may be related to environmental exposures, it is also possible that clustering of residence at the time of birth or menarche may be due to clustering of other socioeconomic or demographic factors. Evaluation of the contribution of socioeconomic status to clustering of residences at birth and menarche is of special interest. There may be other factors associated with residence not measured in this study. The findings are still of interest for further study in order to understand what those exposures might be. We are now investigating the relation between spatio-temporal clustering of residences and exposures to environmental compounds, such as PAHs and benzene, to provide epidemiologic evidence of this finding.

Since the publication of John Snow's [26] well-known cholera map for the city of London in the 19th century, the relationship between the environment and disease has been one of the major research themes in medical geography. Geographic perspectives are of great use in describing geographical patterns of diseases, generating hypotheses on disease etiology, monitoring high risk areas of disease incidence, and suggesting possible causal factors of particular disease [27, 28]. Our study demonstrated that these GIS-based clustering analyses provide effective ways to explore spatial–temporal patterns of clustering. The findings show consistent results; the cluster identified by spatial analyses

remained significant when traditional epidemiologic methods were used, and it was not explained by potential confounders. A recent study comparing 'traditional' epidemiological methods, GIS, and point pattern analysis for use in the spatially referenced public health data concluded that results complement, rather than contradict or duplicate each other [29].

In summary, this analysis of breast cancer clustering in space provides evidence of geographic clustering of pre-menopausal, but not post-menopausal, breast cancer cases at the time of birth and menarche, suggesting a possible influence of exogenous risk factors on breast cancer at these time points. While it is not clear from these data what caused this spatial clustering, it is provocative in providing evidence of the importance of this early period in breast carcinogenesis. Further investigations on genetic susceptibility may be of relevance to identify different effects on pre- and post-menopausal breast cancer. It will also be meaningful to see whether there is temporal clustering of early-life residences as well as spatial clustering. This type of study also needs to be replicated in other settings.

Acknowledgements

This work was supported in part by NIH Grants 1R01 ES09816-01, 1R21 CA87138-01, and U.S. Army Medical Research Grants DAMD17-03-1-0475, DAMD17-00-1-0417.

References

1. Sturgeon SR, Schairer CG, McAdams M, Brinton LA, Hoover RN (1995) Geographic variation in mortality from breast cancer among white women in the United States. *J Natl Cancer Inst* **87**: 1846–1853.
2. Lacey JV, Jr., Devesa SS, Brinton LA (2002) Recent trends in breast cancer incidence and mortality. *Environ Mol Mutagen* **39**: 82–88.
3. Wolff MS, Collman GW, Barrett C, Huff J (1996) Breast cancer and environmental risk factors: epidemiologic and experimental findings. *Annu Rev Pharmacol Toxicol* **36**: 573–596.
4. Laden F, Hunter DJ (1998) Environmental risk factors and female breast cancer. *Annu Rev Public Health* **19**: 101–123.
5. Kuh D, Ben-Shlomo Y (1997) A Life Course Approach to Chronic Disease Epidemiology: Tracing the Origins of Ill-health from Early to Later Life. Oxford: Oxford University Press.
6. Barker DJP (1992) *Fetal and Infant Origins of Adult Disease*. London: BMJ Publishing.
7. Colditz GA, Frazier AL (1995) Models of breast cancer show that risk is set by events of early life: prevention efforts must shift focus. *Cancer Epidemiol Biomarkers Prev* **4**: 567–571.
8. Trichopoulos D (1990) Hypothesis: does breast cancer originate *in-utero*? *Lancet* **335**: 939–940.
9. Potischman N, Troisi R (1999) *In-utero* and early life exposures in relation to risk of breast cancer. *Cancer Causes Control* **10**: 561–573.
10. Lumey LH (1998) Prenatal oestrogens and breast cancer. *Paediatr Perinat Epidemiol* **12**: 361–365.
11. Elliott P, Wakefield JC, Best NG, Briggs DJ (2000) *Spatial Epidemiology*: *Methods and Applications*. Oxford: Oxford University Press.
12. GDT (2001) GDT Dynamp/2000 Version 11.1, Geographic Data Technology, INC. USA.
13. Ripley BD (1981) *Spatial Statistics*. New York: Wiley.
14. Ripley BD (1977) Modelling spatial patterns. *J R Stat Soc Ser* B **39**: 172–212.
15. Rowlingson B, Diggle PJ (1993) SPLANCS: spatial point pattern analysis code in S-plus. *Comput Geosci* **19**: 627–655.
16. Diggle PJ, Chetwynd AD (1991) Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics* **47**: 1155–1163.
17. Kulldorff M, Nagarwalla N (1995) Spatial disease clusters: detection and inference. *Stat Med* **14**: 799–810.
18. Kulldorff M (1997) A spatial scan statistic. *Commun Stat-Theor M* **26**:1481–1496.
19. Bailey TC, Gatrell AC (1995) *Interactive Spatial Data Analysis*. Harlow: Longman.
20. Timander LM, McLafferty S (1998) Breast cancer in West Islip, NY: a spatial clustering analysis with covariates. *Soc Sci Med* **46**: 1623–1636.
21. Rogerson PA, Han D (2002) The effects of migration on the detection of geographic differences in disease risk. *Soc Sci Med* **55**: 1817–1828.
22. McPherson K, Steel CM, Dixon JM (2000) ABC of breast diseases: breast cancer epidemiology, risk factors, and genetics. *Br Med J* **321**: 624–628.
23. Stavola B, Hardy R, Kuh D, Sllva I, Wadsworth M, Swerdlow A (2000) Birthweight, childhood growth and risk of breast cancer in a British cohort. *Br J Cancer* **83**: 964–968.
24. Cressie N (1993) *Statistics for Spatial Data*. New York: Wiley.
25. Reader S (2000) Using survival analysis to study spatial point patterns in geographical epidemiology. *Soc Sci Med* **50**: 985–1000.
26. Snow J (1855) On the mode of communication of cholera. In: Snow on cholera being a reprint of two papers by John Snow, M.D. 1936. New York: The Commonwealth Fund.
27. Jones K, Moon G (1987) *Health, Disease, and Society*. London: Routledge.
28. Meade MS, Florin JW, Gesler WM (1988) *Medical Geography*. New York: Guilford.
29. Dunn C, Kingham S, Rowlingson B et al. (2001) Analysing spatially referenced public health data: a comparison of three methodological approaches. *Health Place* **7**: 1–12.

International Journal of Health Geographics

Research

Assessing spatio-temporal variability of risk surfaces using residential history data in a case control study of breast cancer

Daikwon Han*[1,2], Peter A Rogerson[2,3], Matthew R Bonner[1], Jing Nie[1], John E Vena[4], Paola Muti[1], Maurizio Trevisan[1] and Jo L Freudenheim[1]

Address: [1]Department of Social and Preventive Medicine, University at Buffalo, Buffalo, NY 14214 USA, [2]Department of Geography and National Center for Geographic Information and Analysis, University at Buffalo, Buffalo, NY 14261 USA, [3]Department of Biostatistics, University at Buffalo, Buffalo, NY 14214 USA and [4]Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC 29208 USA

Email: Daikwon Han* - dhan@buffalo.edu; Peter A Rogerson - rogerson@buffalo.edu; Matthew R Bonner - bonnerm@mail.nih.gov; Jing Nie - jingnie@buffalo.edu; John E Vena - jvena@gwm.sc.edu; Paola Muti - muti@buffalo.edu; Maurizio Trevisan - trevisan@buffalo.edu; Jo L Freudenheim - jfreuden@buffalo.edu

* Corresponding author

Published: 12 April 2005

International Journal of Health Geographics 2005, **4**:9 doi:10.1186/1476-072X-4-9

Received: 28 February 2005
Accepted: 12 April 2005

This article is available from: http://www.ij-healthgeographics.com/content/4/1/9

Abstract

Background: Most analyses of spatial clustering of disease have been based on either residence at the time of diagnosis or current residence. An underlying assumption in these analyses is that residence can be used as a proxy for environmental exposure. However, exposures earlier in life and not just those in the most recent period may be of significance. In breast cancer, there is accumulating evidence that early life exposures may contribute to risk. We explored spatio-temporal patterns of risk surfaces using data on lifetime residential history in a case control study of breast cancer, and identified elevated areas of risk and areas potentially having more exposure opportunities, defined as risk surfaces in this study. This approach may be more relevant in understanding the environmental etiology of breast cancer, since lifetime cumulative exposures or exposures at critical times may be more strongly associated with risk for breast cancer than exposures from the recent period.

Results: A GIS-based exploratory spatial analysis was applied, and spatio-temporal variability of those risk surfaces was evaluated using the standardized difference in density surfaces between cases and controls. The significance of the resulting risk surfaces was tested and reported as *p*-values. These surfaces were compared for premenopausal and postmenopausal women, and were obtained for each decade, from the 1940s to 1990s. We found strong evidence of clustering of lifetime residence for premenopausal women (for cases relative to controls), and a less strong suggestion of such clustering for postmenopausal women, and identified a substantial degree of temporal variability of the risk surfaces.

Conclusion: We were able to pinpoint geographic areas with higher risk through exploratory spatial analyses, and to assess temporal variability of the risk surfaces, thus providing a working hypothesis on breast cancer and environmental exposures. Geographic areas with higher case densities need further epidemiologic investigation for potential relationships between lifetime environmental exposures and breast cancer risk. Examination of lifetime residential history provided additional information on geographic areas associated with higher risk; limiting exploration of chronic disease clustering to current residence may neglect important relationships between location and disease.

Background

In a recent analysis of breast cancer by New York State's Department of Health, a breast cancer cluster in the Western New York area was identified [1]. One of the objectives of such disease mapping is to generate hypotheses by identifying spatial patterns so that causal processes may be evaluated further by more rigorous epidemiologic study. Spatial analyses have played a valuable role in explaining different health outcomes and in uncovering environmental causes of disease [2,3]. Residential locations at the time of diagnosis have generally been used in these exploratory spatial analyses [4,5]. Disease mapping has been increasingly used to identify spatial patterns with the aid of Geographic Information Systems (GIS) and exploratory spatial analysis tools, and has been a valuable tool for studies of geographic and environmental epidemiology, especially when the causes of disease and their determinant processes are unknown [6-8]. In particular, there has been recent interest in the use of kernel density estimation methods in epidemiologic studies. Density estimation methods have been used to smooth out noise based on functions of the data in surrounding areas and to overcome problems associated with traditional disease mapping [9,10].

Previous studies using exploratory spatial analyses, however, have been based on either residence at the time of diagnosis or current residence, and only a few recent studies have examined disease risk using information on lifetime residence [11,12]. For chronic disease, there is increasing evidence that lifetime exposures may be more relevant in understanding disease etiology. For breast cancer in particular, several of the well established risk factors (age at menarche, age at first birth) are from early life. There is now evidence that childhood and even *in utero* exposures may affect risk [13]. To examine disease clustering, lifetime cumulative exposures or exposures at critical times in a life course may be more strongly associated with risk for breast cancer than exposures from any one time period, especially the recent period.

In this study, we explored spatio-temporal patterns of risk surfaces using data on lifetime residential history in a case control study of breast cancer. We had previously identified geographic clustering of residence at critical points in early life in relation to breast cancer risk [14]. Here we focused on lifetime cumulative exposure in relation to the disease risk. Risk surfaces were created based on the relative densities of cases and controls – this indicated areas with higher case density as being areas with higher breast cancer risk, thus identifying areas potentially having more exposure opportunities. We used residence as a proxy for potential environmental exposures, conducted exploratory spatial analyses of breast cancer, and produced risk surface maps using information on lifetime residence to identify areas with high breast cancer incidence. In particular, we assessed spatio-temporal variability of risk surfaces using the standardized difference in case and control density, and evaluated the potential use of different kernel density estimation methods in applying them to epidemiologic data.

Results

Descriptive characteristics of study participants by menopausal status are presented in Table 1, and characteristics of lifetime residential history for breast cancer cases and controls are summarized in Table 2. One-fourth of the study participants had at least one previous residence outside the study area, and these were excluded from the analysis. For those residences in the study area, we found that cases were somewhat more mobile, averaging 5.8 and 5.4 residences for pre- and postmenopausal participating cases, compared to 4.9 and 5 residences for pre- and postmenopausal participating controls, respectively.

Table 1: Descriptive characteristics of study participants (Mean ± Standard Deviation): WEB Study, 1996–2001.

	Premenopausal		Postmenopausal	
	Case (n = 325)	Control (n = 610)	Case (n = 841)	Control (n = 1495)
Age (years)	44.9 ± 4.6	44.1 ± 4.6	63.0 ± 8.5	63.4 ± 8.9
Education (years)	14.0 ± 2.3	14.2 ± 2.2	13.3 ± 2.6	13.0 ± 2.3
Parity	1.9 ± 1.3	2.0 ± 1.3	3.0 ± 1.9	2.5 ± 1.8
Age at menarche (years)	12.5 ± 1.6	12.6 ± 1.6	12.6 ± 1.6	12.8 ± 1.7
Age at first birth (years)	25.0 ± 5.1	25.8 ± 4.8	23.8 ± 4.7	23.5 ± 4.3
Recent BMI (kg/m^2)	26.7 ± 6.6	27.2 ± 6.8	28.9 ± 6.1	28.4 ± 6.4
Benign breast disease (yes)	37%	21%	33%	22%
Relative with breast cancer (yes)	21%	10%	20%	14%

Table 2: Descriptive characteristics of lifetime residential history for breast cancer cases and controls: WEB Study, 1996–2001;

	Erie and Niagara (n = 15487)		Outside (n = 4752)		Total (n = 20240)	
	Case	Control	Case	Control	Case	Control
			Premenopausal			
Total numbers of residences in lifetime	1661	2767	432	948	2093	3715
Average numbers of residences per participant	5.8	4.9	3.2	3.4	5.0	4.4
Average years in each residence* (Mean ± SD)	5.6 ± 6.0	6.2 ± 6.6	4.3 ± 5.4	3.9 ± 4.9	5.3 ± 5.9	5.5 ± 6.3
			Postmenopausal			
Total numbers of residences in lifetime	4217	6842	1290	2082	5508	8924
Average numbers of residences per participants	5.4	5.0	3.5	3.2	4.8	4.4
Average years in each residence* (Mean ± SD)	6.9 ± 7.3	7.3 ± 7.9	4.8 ± 5.6	5.2 ± 6.3	6.3 ± 7.0	6.7 ± 7.6

* Excludes residences with missing data for length of residence.

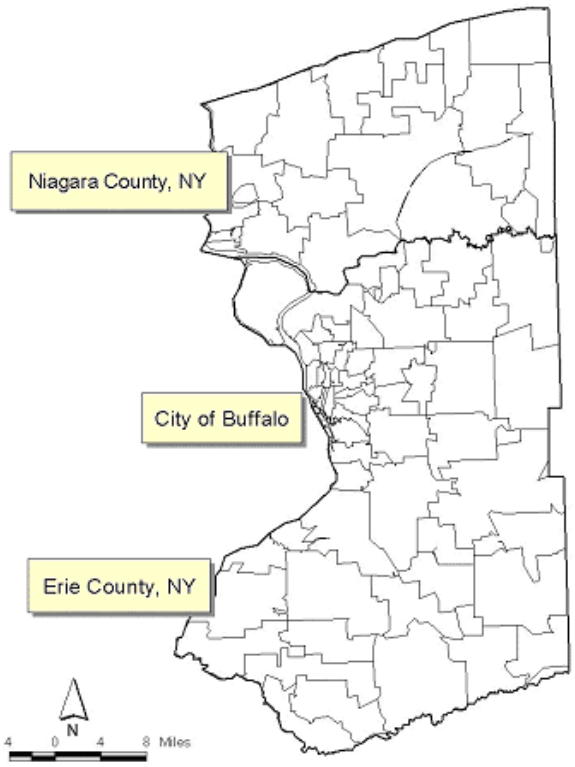

Figure 1
Map of study area: Erie and Niagara counties with zip-code boundaries.

We constructed lifetime cumulative risk surfaces to represent exposure opportunities in lifetime because cumulative exposures may be a more accurate indicator of potential environmental exposures related to breast cancer risk. Figure 1 shows a boundary map of Erie and Niagara counties, while Figure 2 depicts geographic patterns of lifetime residential locations for breast cancer cases and controls by menopausal status. We used the rectangular region as an approximate boundary of the study area to protect individuals' confidentiality. In the study area, there are 4,812 lifetime residential locations for cases (1,328 pre-menopausal and 3,484 post-menopausal residences), and there are 7,886 lifetime residential locations for controls (2,270 pre-menopausal and 5,616 post-menopausal residences).

We evaluated spatial patterns of risk surfaces based on the geographic distribution of lifetime residences in Figure 2. Risk surfaces based on the standardized difference between case and control densities were obtained, and areas with relatively higher case density were identified by menopausal status in Figure 3. In the figure, areas with difference greater than 2 standard deviations (SD) were portrayed as contours and areas exceeding critical values were portrayed as red images. Testing for significance was performed and reported as *p*-values. Those areas with standardized difference greater than 2 SD were quite different between pre- and postmenopausal breast cancer, although ranges of standardized the difference were similar; -6.37 to 4.43 for pre-menopausal, and -6.57 to 3.39 for postmenopausal breast cancer, respectively. There were about 29 rectangular grids in those areas greater than 2 SD for premenopausal, while about 59 grids for postmenopausal breast cancer. Further, the statistical significance of areas must be assessed in light of the fact that multiple areas are tested; these are statistically significant if the difference in density exceeds the critical value of 3.56 at $\alpha = 0.05$ (determined by simulation, where random labelling of cases and controls is carried out). There is one small geographic area of special interest for premenopausal residences in the central and upper region of the city. When these are compared with the geographic

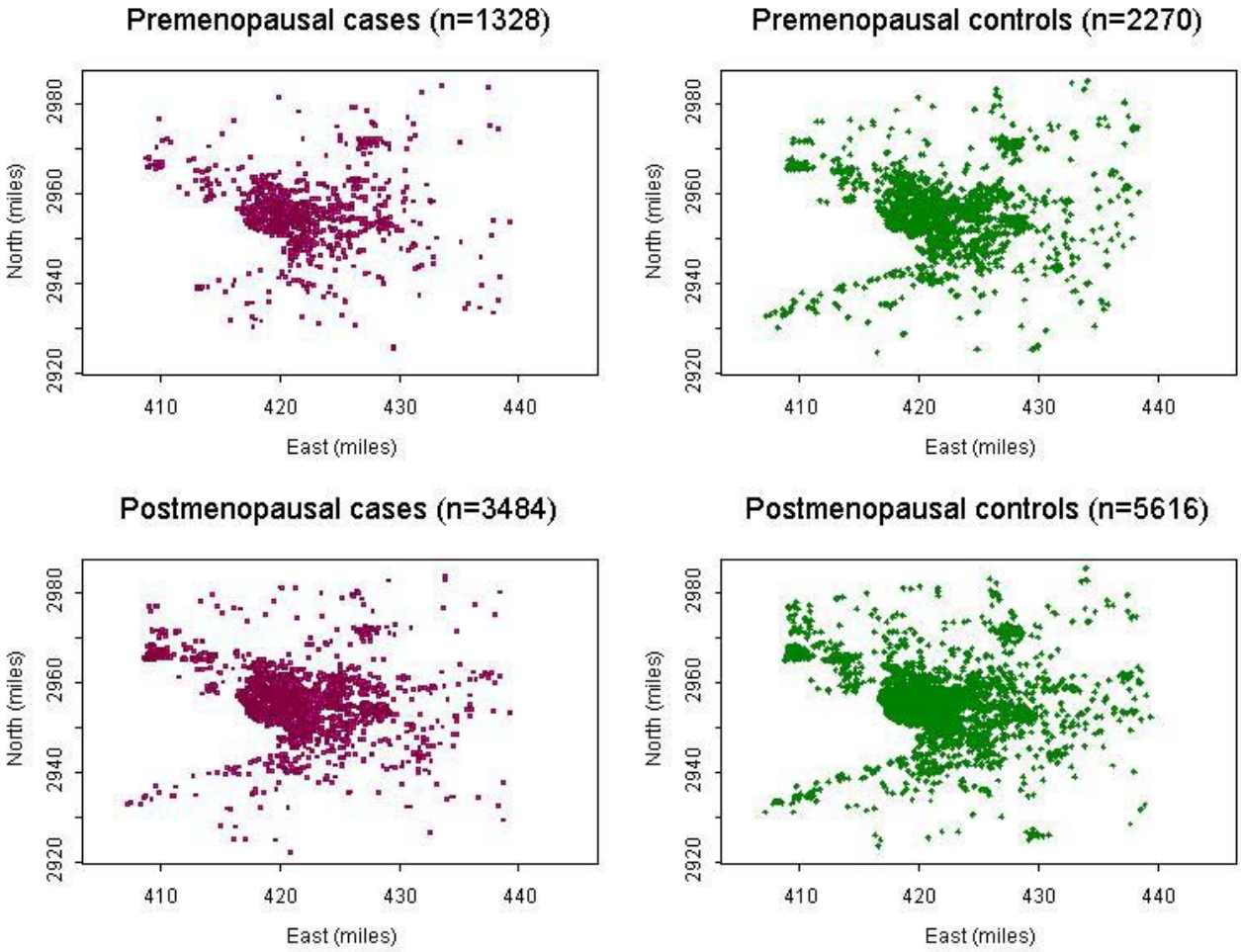

Figure 2
Geographic distribution of breast cancer in Western New York; Shown are all residential locations of breast cancer cases and controls by menopausal status included in the analysis. One dot indicates each residential location. The rectangular region was used as an approximate boundary of the study area instead of actual county boundary in Figure 1. East (x) and north (y) coordinates in projected Universal Traverse Mercator (UTM) miles.

location of clusters of birth and menarche residences identified previously [14], we found that the size and location of these areas are about the same as clusters of menarche residences, but somewhat smaller than the clusters of birth residences. Thus, it is more likely that the same individuals are in both clusters. For post-menopausal residences, no area exceeding the critical value was detected.

Next, we evaluated effects of other risk factors on the risk surfaces. To create age-adjusted risk surfaces, the standardized difference between case and control densities stratified by menopausal status and age groups was examined. Table 3 presents the variability of risk surfaces when stratified by menopausal status and age groups. While there were similar numbers of geographic areas greater than 2 SD regardless of menopausal status and age groups, we found areas (about 8 rectangular grids) greater than critical values only for residences for premenopausal women, aged 35–44, and the geographic location of those areas was identical to the areas identified in Figure 3. In addition, we evaluated the effects of one known risk factor, nulliparity, on those spatial patterns of risk surfaces. Risk surfaces were examined in two groups, nulliparious women and those with at least one child. We observed no difference in spatial patterns of risk surfaces for these two

Table 3: Standardized difference of residences for premenopausal and postmenopausal breast cancer cases and controls by age groups;

Premenopausal	Case (n = 1328)	Control (n = 2270)	Standardized differences	No. of areas[a] > 2SD	No. of areas[a] > critical values[b]
Ages 35–44 (n = 1889)	647	1242	-5.12–5.88	34	8
Ages 45–56 (n = 1709)	681	1028	-4.72–3.14	30	0
Postmenopausal	**Case (n = 3484)**	**Control (n = 5616)**	**Standardized differences**	**No. of areas > 2SD**	**No. of areas > critical values**
Ages 40–64 (n = 4916)	2115	2801	-3.52–3.77	49	0
Ages 65–79 (n = 4184)	1369	2815	-4.50–3.15	40	0

[a] Areas refer to the rectangular grid overlaid on to the study area, [b] Critical values of 3.88 for premenopausal and 3.71 for postmenopausal residences.

Table 4: Standardized difference of residences for premenopausal and postmenopausal breast cancer cases and controls by decades;

	Decades	Case	Control	Standardized differences	No. of areas[a] > 2SD	No. of areas[a] > critical values[b]
Premenopausal	1940s	36	86	-2.35–2.12	2	0
	1950s	222	461	-2.45–3.15	9	0
	1960s	341	633	-4.02–3.78	21	2
	1970s	514	1025	-3.89–3.96	35	2
	1980s	552	1005	-4.26–4.50	24	3
	1990s	457	723	-4.06–4.14	32	5
Postmenopausal	1930s	389	676	-2.51–2.62	24	0
	1940s	809	1253	-3.01–3.04	23	0
	1950s	1158	1968	-3.66–2.91	22	0
	1960s	1247	2072	-4.18–2.99	18	0
	1970s	1100	1839	-4.98–3.13	21	0
	1980s	960	1651	-3.84–2.85	23	0
	1990s	943	1542	-4.40–3.57	22	0

[a] Areas refer to the rectangular grid overlaid on to the study area, [b] Critical values of 3.88 for premenopausal and 3.71 for postmenopausal residences.

groups.; for premenopausal women, there were seven and two geographic areas greater than critical values for nulliparious and parous women, respectively, but none was detected for either group of postmenopausal residences (data not shown).

Lastly, we were interested in evaluating temporal variability of risk surfaces; the standardized difference was obtained for each decade, for both pre- and post-menopausal residences, from the 1940s to 1990s (Table 4). While there was not much difference in the number of areas greater than 2 SD for both menopausal groups, we were able to find geographic areas greater than critical values in the 1960s through 1990s only for residences of premenopausal women. This is consistent with results from the above spatial analysis.

Discussion and conclusions

This study explored the use of kernel density estimation methods to identify spatio-temporal patterns of risk surfaces in a case-control study of breast cancer. We used standardized differences between case and control densities to produce risk surfaces. These risk surfaces were assessed for both pre- and postmenopausal breast cancer.

We found a general tendency for spatial clustering of breast cancer cases, and observed stronger evidence of geographic clustering for pre-menopausal women than for postmenopausal women. Geographic areas greater than 2 SD of the standardized difference were identified among lifetime residences for pre- and postmenopausal women, but more rigorous testing showed such evidence only for premenopausal residences. We were able to pinpoint geographic areas with relatively higher case densities, and to assess temporal variability of risk surfaces.

This study focused on the investigation of breast cancer risk associated with lifetime residential history using GIS-based exploratory spatial analyses. Since environmental risk factors are of continuing interest in breast carcinogenesis, this approach may be more relevant in understanding the environmental etiology of breast cancer [15,16]. Residential location has often been used as a proxy for exposures, and the relationships between residential environment and breast cancer risk have been a focus in recent epidemiologic studies [17-19]. Although the role of clustering analyses remains controversial in scientific advances in our understanding of disease etiology [20,21], these GIS-based exploratory spatial analyses are well suited for environmental epidemiologic investigations; this study demonstrated that smoothed risk surfaces created by kernel methods are useful for large sets of data in space and time, but also when the form of cluster is not well defined. This method can be applied to other epidemiologic analyses. For example, this GIS-based spatial analysis can be effectively used in exposure analyses and assessment, as previously used in identifying people potentially exposed to environmental risk factors [22,23].

Given that these are exploratory methods, however, it is meaningful to compare the strengths and limitations of different approaches when applied to epidemiologic data. We have chosen the standardized difference approach to represent risk surfaces, as opposed to other methods (such as risk ratios) that can be used to create risk surfaces, because it is more easily able to handle the difficulties that arise with small densities. Unsmoothed risk surfaces are relatively easy to manipulate, but they are sensitive to geographic scales, while the ratio of case to control density results in unstable risk surfaces due to small number problems. We were able to reduce this problem in creating risk surfaces based on a standardized difference approach. The selection of optimal bandwidths in the application of kernel methods to cluster detection, and comparison of different types of bandwidths, such as adaptive kernel, will be a subject of future study [24].

It is important to note that current approaches to obtaining density surfaces of lifetime breast cancer risk are limited in several ways. First, we are using residential locations of breast cancer cases and controls; we obtained the difference in densities (risk surfaces) based on their residential location to identify areas with relatively higher case density. We do not know actual exposures and breast cancer risk associated with residential locations. However, this study provides evidence of differential exposure opportunities among cases and controls for further epidemiologic assessment, since spatio-temporal clustering of residential locations in a life course may be an indicator of differential exposure opportunities and the subsequent risk of breast cancer. Another limitation of this analysis is that we were unable to incorporate length of residence into the model. We were able to visualize risk surfaces with different weights based on the actual length of residence of each individual, but there were difficulties in incorporating this information into risk surfaces due to a substantial degree of variability in the length of residence. However, we observed spatial patterns consistent with those in Figure 3, despite the greater variability, when we visualized risk surfaces with length of residence information.

There is also need for cautious interpretation of these results due to the potential for selection bias inherent in the study design, including factors such as non-participation, and missing and excluded residential location. Although we had a relatively stable population and about 40% of study participants were lifetime residents in the study area, there may be different geographic patterns among those included and excluded groups. We excluded missing residential information and residential locations outside of the study area. In addition, we had a rate of non-participation of about 30% in the survey. However, in our earlier study [14], we found that characteristics of subjects in the study area were not different from characteristics of subjects with missing residential information and from subjects excluded due to residence outside of the study area, and that the geographic distribution of participants was not different from that of non-participants. Although we used the same methods for both cases and controls, and interviewers were blinded as to case and control status, there may be recall bias in the lifetime self-reported residential history. We validated the accuracy and reliability of lifetime residential history, especially earlier residences. Our finding was that reported residence information was generally correct. The greater problem was missing data; for this we conducted searches of historical records, as we have described earlier. We are now in the process of obtaining birth addresses from birth certificates for additional validation.

This study has significant implications for further studies on environmental exposure and breast cancer. We had previously found strong evidence of clustering of residence in early life, especially residence at birth and

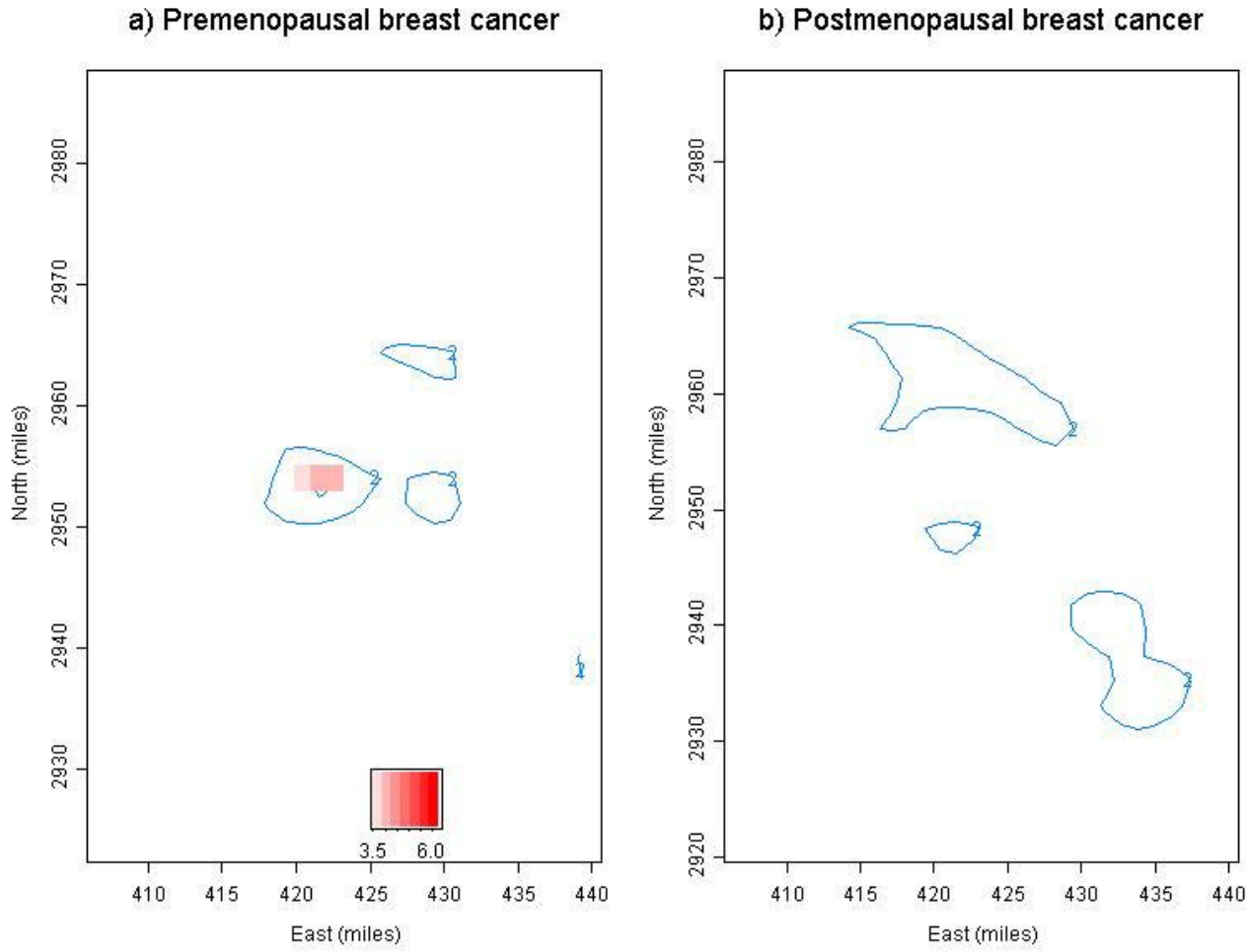

Figure 3
Risk surfaces of pre- and postmenopausal breast cancer using standardized difference; Areas with standardized difference greater than 2 SD are portrayed as contours of 2, and areas exceeding critical value of 3.56 as red images. The rectangular region was used as an approximate boundary of the study area instead of actual county boundary in Figure 1. East (x) and north (y) coordinates in projected UTM miles.

menarche [14]. In these analyses, we were able to show evidence of clustering of lifetime residence. In addition, we found breast cancer cases were more mobile than controls, and that premenopausal participants were more mobile than postmenopausal pariticipants. Average years at current residence was between 11 and 12 years for premenopausal participating cases and controls, compared to 22 and 23 years for postmenopausal participating cases and controls, respectively. Taking findings from this and our previous study together, it appears that examination of exposure opportunities in the past and across the lifespan may be critical for understanding environmental exposures related to breast cancer. Exploring spatio-temporal patterns of lifetime residential history may provide a link between this potential exposure and breast cancer risk. This study provides a more comprehensive analytical framework for the analysis of environmental exposures in relation to breast cancer by considering these components, such as migration and latency periods, and these spatio-temporal patterns of lifetime residential history may be a key to the understanding the actual relationships between environmental exposure and subsequent breast cancer risk.

To provide more accurate measures of personal cumulative exposures based on complete lifetime residential

history, further studies should take into account the different effects of time periods or timing of exposures; construction of lifetime cumulative risk surfaces with different weights for exposure sensitivity at different points in time is a potential improvement and explanation of the methods employed to date. This approach may help to provide an answer to the question of where, when, and what kinds of exposures have influenced individuals' risk for a particular disease. Further, there has been recent development of a GIS-based framework to examine spatio-temporal patterns of lifetime residential history; the geospatial lifeline concept and space-time information system (STIS) approach is a good example of this [25-27]. We are currently testing the feasibility of similarity and difference measures of an individual's lifetime residential history using this case and control data, since it would be a powerful tool to analyze the personal environmental exposures associated with lifetime residential history.

In summary, we found evidence of clustering of lifetime residence for premenopausal cases relative to controls, and a substantial degree of spatio-temporal variability in the risk surfaces, thus providing a working hypothesis on breast cancer and environmental exposures. Geographic areas with higher case densities need further epidemiologic investigation for potential relationships between lifetime environmental exposures and breast cancer risk. Examination of lifetime residential history provided additional information on geographic areas associated with higher risk; limiting exploration of chronic disease clustering to current residence may neglect important relationships between location and disease. Further studies on the relationship between disease risk and environmental exposures associated with lifetime residential history should be replicated in other settings.

Methods

The Western New York Exposure and Breast Cancer Study (WEB Study)

Data from a population-based case-control study of breast cancer in western New York (the WEB Study) were used for our analyses. Participants were women, age 35–79 who were residents of Erie and Niagara counties, with no history of cancer other than non-melanoma skin cancer; cases were women with incident, primary, pathologically confirmed breast cancer, diagnosed during the period 1996–2001, and controls were randomly selected and frequency matched to cases on age, race, and county of current residence. Details of the WEB study, including selection, ascertainment, in-depth interview processes, have been described previously [14]. We collected lifetime residential histories for 1,166 cases and 2,105 controls, identified 20,240 lifetime addresses, an average of approximately 6 addresses for each individual, from participating cases and controls. Analyses were restricted to those residential locations within the two counties of study area.

Geocoding of residential location

Geocoding of residential locations enables us to record each individual's locational information as x and y coordinates to be used in further spatial analyses. Address geocoding is a process that creates a theme based on the address data in a tabular form (event theme) and a reference feature theme (street map) to add point locations defined by the street address to the map. Matching depends not only on the quality of the reference theme, but also on the quality of the tabular data to be mapped. We used GDT/Dynamap 2000, an enhanced version of Topographically Integrated Geographic Encoding and Referencing system (TIGER), as a reference theme. In a study validating the positional accuracy of TIGER for the use in epidemiologic study [28], we found positional accuracy to be extremely high.

The overall matching rate for the 15,487 Erie and Niagara County residential locations was 82% (12,698); 91% of the matches were matched with complete information (good match), while about 9% of were estimated with partial information (interactive match) as detailed below. Geocoding success rates were lower for earlier residences, mainly due to more missing and partial information of earlier residences and changes in streets names and zip codes. However, it is important to note that there were few changes in street structure for this region during the time period of interest. There was, of course, addition of new streets, but existing streets were unchanged for the most part, and thus we found that the process of geocoding using current information was not inappropriate. We utilized various resources, including historic city directories with address information for residents, historical maps, and commercial address databases, to find missing residence information, and we developed several strategies to improve matching rates [14]. In addition, for residences where we had a known street name but no known street number and if the total length of the street was one kilometre or less, we estimated the residence location as the midpoint of the street.

Kernel density estimation methods

Kernel density estimation methods have been used for disease mapping and for the detection of geographical location of clusters in epidemiologic settings [10,12,29]. A general form of the kernel k is defined as, $k_{\mathrm{h}}(x) = \frac{1}{h} k(\frac{x}{h})$, and by averaging over these individual kernel functions, we obtain the kernel density estimator, $\hat{f}_{\mathrm{h}}(x)$;

$$\hat{f}_{\mathrm{h}}(x) = \frac{1}{n}\sum_{\mathrm{i}=1}^{\mathrm{n}} k_{\mathrm{h}}(x\text{-}x_{\mathrm{i}}) = \frac{1}{nh}\sum_{i=1}^{n} k_h(\frac{x-x_i}{h})$$

where x is the location for density estimation, x_i is the observed point location, n is the number of points, and h is a smoothing parameter that regulates the degree of smoothness. Kernel functions are symmetric around zero and integrate to one. Both kernel and kernel density estimator are density functions; thus $\int k_{\mathrm{x}}(x)\mathrm{dx} = 1$ and $\int f_{\mathrm{h}}(x)\mathrm{dx} = 1$. For the two-dimensional kernel density estimator:

$$\hat{f}_{\mathrm{h}}(x_1, x_2) = \frac{1}{n}\sum_{\mathrm{i}=1}^{\mathrm{n}} k_{\mathrm{h}}(x_1 - x_{1\mathrm{i}}; h_1) k_h(x_2 - x_{2i}; h_2) \qquad (1)$$

While the estimation of intensity for one point pattern may show patterns of high and low risk areas, this is of limited utility in epidemiologic applications because it may largely reflect the pattern of population distribution [30]. The ratio of case to control density can be more effectively used in epidemiologic settings [9,31]. To estimate the ratio of the two density estimates, the ratio of case to control density, $\frac{\hat{f}(x)}{\hat{g}(x)}$ is used, where $\hat{f}(x)$ and $\hat{g}(x)$ are estimates of the intensity for cases and controls, respectively. Smoothed risk surfaces using the ratio or log ratio of case to control density can be less effective with small sample sizes; it is possible to have spurious, high risk areas in the application of the ratio of densities when the value of the control density is too small.

The relative difference between two densities provides an alternative way to assess the spatial variability of risk surfaces. Using the square root variance stabilizing transformation and the standardized difference, this measure allows us to identify areas with differences between case and control density exceeding two standard deviations [32]. The standardized difference between case and control densities is obtained by taking the square root of the case density minus the square root of the control density, and dividing by the standard deviation of the difference between the densities.

$$\frac{\sqrt{\hat{f}(x)} - \sqrt{\hat{g}(x)}}{\sqrt{\mathrm{var}\left(\sqrt{\hat{f}} - \sqrt{\hat{g}}\right)}} \qquad (2)$$

where $\mathrm{var}\{\sqrt{f} - \sqrt{g}\} \approx \frac{1}{4}\frac{1}{h}(\frac{1}{n_1} + \frac{1}{n_2})\int {k_h}^2(x)dx$

Analytical procedures

The first step in creating smoothed risk surfaces using kernel methods was to create reference grids and overlay the study area with them. We obtained the smoothed intensity for both cases and controls by calculating the distance between each point on the reference grid and the locations of breast cancer cases and controls. We used the quartic kernel to estimate the intensity of points at each grid point, although the choice of kernel type is not crucial as long as the kernel is symmetrical [33]. We applied equal, fixed bandwidths for both cases and controls using equation (1) because the objective is to describe overall patterns of the underlying spatial distribution [9]. The risk surface was obtained by forming the difference in densities based on equation (2).

The above analyses were repeated with varying bandwidths because the choice of appropriate bandwidth is one of the primary concerns of the kernel method. Although a subjective choice made from a range of values is commonly used [33], selections of bandwidth were made here on the basis of several factors. To avoid subjectivity, we first began with the commonly used optimal bandwidth designed to minimize the estimated mean square error [34]. Other established methods, such as cross validation, were tried and these resulted in small bandwidths because of the large sample size [34,35]. In addition, to take into account the spatial distribution of point patterns and to avoid problems associated with fixed bandwidths, we initially selected bandwidths based on the average distance among points. Since the size of the study area is approximately 30 miles in width and 60 miles in length, we selected a one-mile radius as an initial bandwidth of the kernel, with a range of 0.5 to 10 miles. In summary, we searched over a range of bandwidths and ultimately chose a two-mile bandwidth as a balance between over-smoothing and under-smoothing. Increasing bandwidth implies increasing the amount of smoothing in the estimate. A larger bandwidth results in very smooth density surfaces, while too small a bandwidth produces noisy density estimates. These issues are important in applications of our case-control data; the geographic distribution of cases and controls is dependent on the population distribution, which is greatly concentrated in urban areas and is sparser in rural areas. Thus, the use of a small bandwidth less than two miles resulted in an unrecognizable pattern of density.

Risk surfaces based on the ratio and difference in density surfaces between cases and controls were implemented in a GIS and S-Plus environment [36]. In addition to finding the risk surface associated with the standardized difference between densities, we tested the significance of the difference surfaces between cases and controls by Monte Carlo simulation. Under the null hypothesis of constant risk in the study area, we obtained critical values. We first randomly assigned case and control status to each of the case and control locations, based on the proportion of

cases and controls among the set of cases and controls. Then we obtained the 95th percentile for the maximum difference between case and control densities from 999 simulations.

Authors' contributions

All authors have participated fully in the conduct of this study, the acquisition of data, the analysis and/or the preparation of the manuscript in a substantial manner. DH performed the statistical analysis and drafted the manuscript; PAR and JLF provided critical review and input. JLF, JEV, MAB, JN, PM and MT participated in the design of the study, participated in interpretation, as well as in data acquisition efforts. All authors read and approved this manuscript.

Acknowledgements

This work was supported in part by NIH Grants 1R01 ES09816-01, 1R21 CA87138-01, and U.S. Army Medical Research Grants DAMD17-03-1-0475, DAMD17-00-1-0417.

References

1. New York State Department of Health: **The New York State Cancer Surveillance Improvement Initiative.** Albany, NY, ; 1998.
2. Cromley EK, McLafferty SL: **GIS and Public Health.** New York, The Guilford Press; 2002.
3. Khan OA, Skinner R: **Geographic Information Systems and Health Applications.** Hershey, PA, Idea Group Publishing; 2003.
4. Rushton G, Lolonis P: **Exploratory spatial analysis of birth defect rates in an urban population.** *Stat Med* 1996, **15:**717-726.
5. Timander LM, McLafferty S: **Breast cancer in West Islip, NY: a spatial clustering analysis with covariates.** *Soc Sci Med* 1998, **46:**1623-1635.
6. Lawson A, Biggeri A, Bohning D, Lesaffre E, Viel J, Bertollini R: **Disease Mapping and Risk Assessment for Public Health.** Chichester, John Wiley and Sons; 1999.
7. Rushton G: **Improving the geographic basis of health surveillance using GIS.** In *GIS and Health* Edited by: Gatrell, AC and Löytönen, M. London, Taylor and Francis; 1998:63-80.
8. Elliott P, Wakefield JC, Best NG, Briggs DJ: **Spatial Epidemiology: Methods and Applications.** Oxford, Oxford University Press; 2000.
9. Kelsall JE, Diggle PJ: **Non-parametric estimation of spatial variation in relative risk.** *Stat Med* 1995, **14:**2335-2342.
10. Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS: **Spatial point pattern analysis and its application in geographical epidemiology.** *Transactions of the Institute of British Geographers* 1996, **21:**256-274.
11. Paulu C, Aschengrau A, Ozonoff D: **Exploring associations between residential location and breast cancer incidence in a case-control study.** *Environ Health Perspect* 2002, **110:**471-478.
12. Sabel CE, Gatrell AC, Loytonen M, Maasilta P, Jokelainen M: **Modelling exposure opportunities: estimating relative risk for motor neurone disease in Finland.** *Soc Sci Med* 2000, **50:**1121-1137.
13. Potischman N, Troisi R: **In-utero and early life exposures in relation to risk of breast cancer.** *Cancer Causes Control* 1999, **10:**561-573.
14. Han D, Rogerson PA, Nie J, Bonner MR, Vena JE, Vito D, Muti P, Trevisan M, Edge SB, Freudenheim JL: **Geographic clustering of residence in early life and subsequent risk of breast cancer (United States).** *Cancer Causes Control* 2004, **15:**921-929.
15. Brody JG, Rudel RA: **Environmental pollutants and breast cancer.** *Environ Health Perspect* 2003, **111:**1007-1019.
16. DeBruin LS, Josephy PD: **Perspectives on the chemical etiology of breast cancer.** *Environ Health Perspect* 2002, **110 Suppl 1:**119-128.
17. Gammon MD, Neugut AI, Santella RM, Teitelbaum SL, Britton JA, Terry MB, Eng SM, Wolff MS, Stellman SD, Kabat GC, Levin B, Bradlow HL, Hatch M, Beyea J, Camann D, Trent M, Senie RT, Garbowski GC, Maffeo C, Montalvan P, Berkowitz GS, Kemeny M, Citron M, Schnabe F, Schuss A, Hajdu S, Vincguerra V, Collman GW, Obrams GI: **The Long Island Breast Cancer Study Project: description of a multi-institutional collaboration to identify environmental risk factors for breast cancer.** *Breast Cancer Res Treat* 2002, **74:**235-254.
18. Lewis-Michl EL, Melius JM, Kallenbach LR, Ju CL, Talbot TO, Orr MF, Lauridsen PE: **Breast cancer risk and residence near industry or traffic in Nassau and Suffolk Counties, Long Island, New York.** *Arch Environ Health* 1996, **51:**255-265.
19. McKelvey W, Brody JG, Aschengrau A, Swartz CH: **Association between residence on Cape Cod, Massachusetts, and breast cancer.** *Ann Epidemiol* 2004, **14:**89-94.
20. Rothman KJ: **A sobering start for the cluster busters' conference.** *Am J Epidemiol* 1990, **132:**S6-13.
21. Wartenberg D: **Investigating disease clusters: why, when and how ?** *Journal of the Royal Statistical Society A* 2001, **164:**13-22.
22. Ward MH, Nuckols JR, Weigel SJ, Maxwell SK, Cantor KP, Miller RS: **Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System.** *Environ Health Perspect* 2000, **108:**5-12.
23. Wartenberg D, Greenberg M, Lathrop R: **Identification and characterization of populations living near high-voltage transmission lines: a pilot study.** *Environ Health Perspect* 1993, **101:**626-632.
24. Brunsdon C: **Estimating probability surfaces for geographical point data: an adaptive kernel algorithm.** *Computers and Geosciences* 1995, **21:**877-894.
25. Jacquez GM, Greiling DA, Kaufmann AM: **Design and implementation of a space-time intelligence system for disease surveillance.** *Journal of Geographical Systems* 2005, **7:**7-23.
26. Rushton G, Elmes G, McMaster R: **Considerations for improving geographic information system research in public health.** *URISA Journal* 2000, **12:**31-49.
27. Mark DM, Egenhofer MJ, Bian L, Hornsby KE, Rogerson PA, Vena JE: **Spatiotemporal GIS analysis for environmental health: solutions using geospatial lifeline.** In *Geography and medicine: GEOMED '99* Edited by: Flahaut A, Toubiana L and Valleron AJ. Paris, Elsevier; 2000:65-78.
28. Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Freudenheim JL: **Positional accuracy of geocoded addresses in epidemiologic research.** *Epidemiology* 2003, **14:**408-412.
29. Bithell JF: **An application of density estimation to geographical epidemiology.** *Stat Med* 1990, **9:**691-701.
30. Bailey TC, Gatrell AC: **Interactive Spatial Data Analysis.** Harlow, England, Longman Scientific & Technical; 1995.
31. Kelsall JE, Diggle PJ: **Spatial variation in risk of disease: a nonparametric binary regression approach.** *Applied Statistics* 1998, **47:**559-573.
32. Bowman AW, Azzalini A: **Applied Smoothing Techniques for Data Analysis.** Oxford, Clarendon Press; 1997.
33. Silverman BW: **Density Estimation for Statistics and Data Analysis.** London, Chapman and Hall; 1986.
34. Diggle PJ: **Statistical Analysis of Spatial Point Patterns.** London, Academic Press; 1983.
35. Bowman AW: **An alternative method of cross-validation for the smoothing of density estimates.** *Biometrika* 1984, **71:**353-360.
36. Venables WN, Ripley BD: **Modern Applied Statistics with S-Plus.** New York, Springer; 1998.

Lifetime Adult Weight Gain, Central Adiposity, and the Risk of Pre- and Postmenopausal Breast Cancer in the Western New York Exposures and Breast Cancer Study

Daikwon Han[1], Jing Nie[1], Matthew R. Bonner[1], Susan E. McCann[2], Paola Muti[3], Maurizio Trevisan[1], Farah Ramirez[4], Dominica Vito[1], Jo L. Freudenheim[1]

Submitted to International Journal of Cancer

[1] Department of Social and Preventive Medicine, University at Buffalo-State University of New York, Buffalo, NY 14214.

[2] Department of Epidemiology, Roswell Park Cancer Institute, Buffalo, NY 14263.

[3] Department of Cancer Epidemiology, Italian National Cancer Institute.

[4] Department of Physical Education and Recreation, University of Puerto Rico.

Corresponding author: Daikwon Han, Institute for Regional Analysis and Public Policy, 201 Lloyd Cassity, Morehead, KY 40351

Phone: (606)783-2300, Fax: (606)783-5092, Email: d.han@moreheadstate.edu

Acknowledgements: This work was supported in part by U.S. Army Medical Research Grants DAMD17-03-1-0475, DAMD17-00-1-0417.

Abstract

While there are quite consistent data regarding associations of body weight and postmenopausal breast cancer, there are now accumulating data that would indicate that weight gain in adult life is more predictive of risk than absolute body weight. There is, however, little known about the relative impact of timing of weight gain in adult life as well as other characteristics of the weight and breast cancer association that might provide insight into the mechanism of the observation. We conducted a population-based case control study of breast cancer (1996-2001), the Western New York Exposures and Breast Cancer Study (WEB Study). Included were 1,166 women with primary, histologically confirmed, incident breast cancer and 2,105 controls frequency-matched on age, race, and county of residence. Unconditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI). We found increased risk of breast cancer associated with lifetime adult weight gain among post- but not premenopausal women. Further, adult weight gain was associated with risk in postmenopausal women only among those in strata defined by waist circumference above median, by positive estrogen or progesterone status, or by never use of hormone replacement therapy. We also found an association with risk for weight gain since first pregnancy and for weight gain between the time of the first pregnancy and menopause, independent of body mass index and lifetime adult weight gain. Our results suggest that there are time periods of weight gain that have greater impact on risk, and that central body fat, receptor status and hormone replacement therapy may all affect the observed association.

Numerous epidemiologic studies of the relationship between body size and breast cancer risk have been conducted to examine its potential role as a modifiable risk factor, independent of dietary intake and physical activity.[1-4] Although our ability to explain the mechanism of the observed association is still limited, there is quite consistent evidence showing an association of indicators of body size and postmenopausal breast cancer. In particular, body mass index (BMI) and central adiposity have been shown to be associated with increased risk of post- but not premenopausal breast cancer.[5-9]

There are now accumulating data that would indicate that weight gain in adult life is more predictive of risk than absolute body weight or BMI.[10-15] There is, however, little known about the timing of weight gain in adult life as well as other characteristics of the weight and breast cancer association that might provide insight into the mechanism of the observation. There is interest in weight gain during particular periods of life, especially weight gain during periods of hormonal changes, such as pregnancy and menopause.[4,16-17] Understanding of timing of weight gain in relation to risk could provide insight into the mechanism of the observed associations. In addition, there is evidence that central adiposity is associated with breast cancer risk[7-9,18-19]; interactions between weight gain and central adiposity are of interest, because of difference in metabolic activity depending on the location of fat disposition.

We examined here associations of lifetime adult weight gain with pre- and postmenopausal breast cancer, examining in particular the role of the timing of weight gain on postmenopausal breast cancer risk. We examined weight change between each decade as well as weight change at specific time points in a woman's life: around the time of women's first pregnancy and at menopause, as important time periods of

hormonal change. Interrelationships between central adiposity and weight gain were also assessed to explain effects of adult weight gain on subsequent risk of postmenopausal breast cancer. Further, we evaluted whether effects of weight gain on postmenopausal breast cancer differed by estrogen and progesterone receptor status of tumors and by a women's use of hormone replacement therapy.

Methods

Study population

We conducted a population-based case control study of breast cancer (1996-2001), the Western New York Exposures and Breast Cancer Study (WEB Study). Eligible for the study as cases were all women diagnosed with primary, histologically-confirmed, incident breast cancer, age 35-79 and resident of the two counties of the study area. All cases were interviewed within one year of diagnosis; most were interviewed within three to six months following diagnosis. Controls were randomly selected from the New York State Department of Motor Vehicles driver's license list (≤ aged 65) and the Health Care Finance Administration rolls (> 65 years), frequency-matched to cases on age, race, and county of residence. A total of 1,638 cases and 3,396 controls met our inclusion criteria of age 35-79, current residence in Erie or Niagara County in New York State, no previous cancer diagnosis other than non-melanoma skin cancer. Response rates were 71% (1,166/1,638) and 62% (2,105/3,396) for cases and controls, respectively, among cases and controls for whom we could determine eligibility. All participants provided informed consent, and the protocol was approved by the Institutional Review Boards of the University at Buffalo and of all the participating hospitals.

Data collection

Extensive in-person interviews and self-administered questionnaires were used to ascertain information on potential confounding factors and anthropometric measures, including lifetime weight. Participants were asked to recall their body weight for each decade of their lives from age 20 to one year before diagnosis for cases and one year before interview for controls. They were also asked to recall the amount of their weight gain during their first pregnancy, and to identify their body shape at the time of menarche from among nine pictograms. Current height, weight, and several measures of central adiposity (abdominal height, waist circumference, hip circumference) were measured by trained interviewers according to a standardized protocol. Waist circumference was measured by placing the tape around the smallest point between the top of iliac crest and the bottom of rib cage; hip circumference was measured by placing the tape around the hips at the biggest circumference point between the iliac crest and the crotch, and abdominal height was measured using a caliper on the participant in a recumbent position. All measures were to the nearest 0.1 cm. Three measurements were initially made for accuracy, and these were repeated until the three readings were all within 0.5 cm of each other. For the analyses reported here, current weight was reported weight one year before interview, and BMI based on measured height during interview and reported weight one year before interview were used. Weight one year before interview and measured weight were highly correlated ($r = 0.91$)

Estrogen receptor (ER) and progesterone receptor (PR) status information was abstracted from pathology reports. Receptor status was determined by either biochemical

or immunoperoxidase assay; ER positive tumors were those with at least an ER concentration values greater than 5 fmol/mg or with greater than 5% of the cells with estrogen receptors. We obtained ER and PR status for 646 cases of the 841 eligible postmenopausal breast cancer cases; for the remaining 23% (195) status was unknown either because women did not give us permission to examine their medical records (17) or because the receptor status of the tumor was not measured.

Data analysis

Lifetime adult weight change was calculated as the difference between reported weight at age 20 and weight one year before interview. Weight change was examined for a number of intervals that are potentially important: from the time of a first pregnancy and from the time of menopause to one year before interview, between the time of first pregnancy and menopause, and during each decade of life from age 20.

All analyses were conducted stratified on menopausal status at diagnosis. Women were considered postmenopausal if their menses had ceased permanently and naturally. Among other women, participants were also considered postmenopausal if any of the following conditions were true: they had had a bilateral oophorectomy, they had had a hysterectomy without removal of the ovaries and they were older than 50, their menses had ceased permanently due to radiation or other medical treatment and they were older than 55.

Unconditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI). The cut points for the categorical analyses were derived from the distribution of controls; quartiles of weight gain were determined as well as another

category of women who had not gained or who had lost weight during the time period of study. All models were adjusted for age at interview, education, age at first birth, age at menarche, previous benign breast disease, family history of breast cancer in a first degree relative, and age at menopause and use of hormone replacement therapy (for postmenopausal women only). *P* for trend was determined by the *p*-value for the coefficient of the continuous exposure variable, while adjusting for covariates and excluding the group of women who had lost weight. Since weight change and BMI were highly correlated, we adjusted for BMI by including residuals of the regression of weight gain on BMI.[20]

Results

While details of the study population have been published previously,[21,22] descriptive characteristics of cases and controls by menopausal status, including several anthropometric measures, were presented in Table 1. Among premenopausal participants, these anthropometric measures were not statistically different between cases and controls, with the exception of BMI at age 20 (p=0.03) and body type at menarche (p=0.04). Also weight gain since age 20 was not statistically different between premenopausal cases and controls. Further we examined premenopausal breast cancer risk associated with BMI and weight change between age 20 and one year ago. As has been found by others, among the pre-menopausal women, BMI and lifetime weight gain were not associated with risk; adjusted OR was 0.75 (95% CI 0.49-1.16, p for trend=0.18) for women in the highest quartile of BMI (>30.9 kg/m^2) compared to the lowest (<22.2 kg/m^2), while adjusted OR was 0.84 (95% CI 0.51-1.38, p for trend=0.12) for women in the highest (> 25 kg)

compared to women who gained between 0 and 6.8 kg. Because the association of risk with weight gain was not seen among premenopausal women, further analyses were restricted to postmenopausal participants only.

Selected anthropometric measures of postmenopausal participants are also shown in Table 1; in general, anthropometric measures in early life, including weight at age 20, body type at menarche, weight gained during first full term pregnancy were not different between cases and controls, while measures in later adult life, height, weight one year ago, measured weight, abdominal height, and waist circumference were statistically different (p values of 0.001, 0.003, 0.009, 0.01, 0.001 respectively). Also, correlations between weight, BMI, and weight change between time periods were examined; for subsequent analyses, we adjusted for BMI residuals to identify independent effects of weight change on breast cancer risk; BMI one year ago was correlated with weight change between age 20 and one year ago (r=0.85 and 0.86 for cases and controls, respectively).

Similarly, we found positive associations of postmenopausal breast cancer risk with most anthropometric measurements made at the time of interview when we examined association of various anthropometric measures in different time periods. Recent indicators of body size, including body weight one year before interview, height, BMI one year before interview were all associated with increased risk of postmenopausal breast cancer; women with relatively higher body weight, height, and BMI had an increased risk of 1.61 (95% CI 1.19-2.20, p=0.02), 1.58 (95% CI 1.20-2.09, p=0.001), 1.57 (95%CI 1.18-2.10, p=0.02), respectively when comparing highest to lowest quartile. We examined waist circumference as a measure of central adiposity; it was associated

with postmenopausal breast cancer (adjusted OR 1.76, 95% CI 1.33-2.32). We also examined abdominal height in relation to risk; this measure of central adiposity was highly correlated with waist circumference (r = 0.85); the association with risk was similar to that for waist circumference (data not shown). Absolute body weight at first pregnancy and menopause were also associated with increased risk of postmenopausal breast cancer; women in the highest quartile of weight had a risk of 1.61 (95% CI 1.08-2.40) and 1.98 (95% CI 1.36-2.88), respectively compared with the women in the lowest. Neither BMI at first pregnancy nor BMI at menopause was associated with increased risk of postmenopausal breast cancer. In addition, neither weight nor BMI at age 20 was associated with increased risk.

Risk associated with weight change between age 20 and one year ago for postmenopausal breast cancer are shown in Table 2. An increased risk of breast cancer was found for postmenopausal women who gained more than 27.3 kg compared with women who gained between 0 and 9.1 kg (adjusted OR 1.71, 95% CI 1.23-2.37; p for trend = 0.05), even after controlling for BMI. Each 5 kg increase in the weight between age 20 and one year ago was associated with a 4% change in risk; OR 1.04 (95%CI 1.03-1.05) per 5kg of weight gain in continuous form between two time points. Weight changes between decades and between several time periods with biological relevance to breast cancer were also examined. We found a positive association of weight change between several decades with postmenopausal breast cancer. There was a significant increase in risk of postmenopausal breast cancer associated with weight gain in the age periods 30-39 (OR 1.72, 95% CI 1.14-2.59), 40-49 (OR 1.92, 95% CI 1.17-3.15), and 60-69 (OR 1.94, 95% CI 1.21-3.10), but not in other decades of life; however, confidence

intervals for all decades overlapped. With adjustment for total lifetime weight change, odds ratios were attenuated and all confidence intervals for weight gain during decades of life included the null.

We also examined weight change around the time of women's first pregnancy and the time of menopause, and found significantly increased breast cancer risk for weight change for weight gain between menopause and the present and for weight gain between first pregnancy and menopause. When weight change between first pregnancy and menopause was examined, we found positive associations; adult weight gain between these two hormonal time periods was most strongly associated with the increased risk of postmenopausal breast cancer (OR 1.91, 95% CI 1.26-2.88, comparing highest to lowest quartile). We found similar associations of risk of postmenopausal breast cancer for weight change from first pregnancy to one year ago (OR 1.70, 95% CI 1.22-2.37), and from menopause to one year ago (OR 1.58, 95% CI 1.09-2.30). Again, confidence intervals for these time periods overlapped and these results overlapped with those for the decade analysis. We further adjusted these ORs for weight gain between age 20 and one year ago; ORs were diminished except for the one for weight change since first pregnancy which remained of similar magnitude. On the other hand, weight change between age 20 and first pregnancy was not associated with the risk. We did not observe an association of reported weight gained during the first pregnancy with breast cancer risk (data not shown). Additionally, breast cancer risk associated with weight change between age 20 and one year ago were analyzed by the time since menopause. We found that adult weight gain was more strongly associated with the increased risk of breast cancer among postmenopausal women with a longer time since menopause; there was a

significant increase in risk associated with weight gain for those women in stratum more than 20 years, and 11-20 years since menopause (OR 2.35, 95% CI 1.24-4.46, p-trend=.05 and OR 2.00, 95% CI 1.10-3.64, p-trend=.01 respectively), but not in less than 10 years (OR 1.25, 95% CI 0.74-2.12, p for trend=0.59).

We also examined the relationships between weight change and central adiposity with breast cancer risk. ORs for risk associated with weight change between age 20 and one year ago, stratified by the categories defined by the median of waist circumference were calculated (Table 3). Adult weight gain and central body fat were correlated (r=0.68); there was more weight gain for both cases and controls among those with more central body fat. Adult weight gain was associated with increased risk of breast cancer only among women with waist circumference above the median. Although a test for interaction was not statistically significant (p=0.42), there was a nonmonotonic increase in the group of women with waist circumference above the median, but there was no association with waist circumference below the median.

Finally, we evaluated whether the association of weight gain with risk was different by estrogen and progesterone receptor status. Adult weight gain was strongly associated with increased risk of postmenopausal breast cancer among ER or PR positive tumors, and statistically significant trends were observed (p=0.001, see Table 4). Weight change was not associated with breast cancer risk among ER or PR negative tumors. We also observed that the risk of postmenopausal breast cancer in relation to adult weight gain was more strongly associated with risk among never- than among ever-users of hormone replacement therapy.

Discussion

We investigated the association of lifetime adult weight gain and pre- and postmenopausal breast cancer, in particular weight changes at different time points in a woman's life. As has been reported previously,[15,23-24] we observed increased risk of breast cancer associated with BMI and lifetime adult weight gain among post- but not premenopausal women. We also found that lifetime adult weight gain was associated with postmenopausal breast cancer risk only among those in strata defined by higher waist circumference (above median), by positive estrogen or progesterone status, or by never use of HRT. While current weight, height, and waist circumference were associated with risk, we did not find any association with risk for measures of body size at the time of menarche, weight at age 20, or weight gain during the first pregnancy. In relation to timing of weight gain, we did find an association of weight gain since first pregnancy, and weight gain between the time of the first pregnancy and menopause, independent of BMI one year ago and lifetime adult weight gain. Weight gain since menopause was also associated with risk. Adult weight gain was more associated with risk for women whose menopause was longer ago than for those who had experienced it more recently.

Timing of weight gain has not generally been addressed adequately in previous studies of weight change and breast cancer, and such timing may provide a clue on etiologic role of adult weigh gain in relation to breast cancer risk. We found positive associations with risk for weight gain during the 30's and 40's, and between the time periods of hormonal changes. However, the mechanism for these findings is not known. It could be that weight gain is an indicator of hormonal environment and that the factors

which lead to increased body weight also increase risk. It may also be that energy balance at these time points is particularly significant. Lastly, it may also be that there is more variability in weight change in these time periods and that affects our ability to detect differences in risk. We also found positive associations of adult weight gain with risk for those with positive estrogen or progesterone status and for never users of HRT, consistent with the findings of others.[13,23-27] These findings also lend credence to the notion that the association of weight gain with risk of postmenopausal breast cancer is related to steroid hormone metabolism.

When we examined the relationships between weight gain and central adiposity on breast cancer risk, we found evidence of increased risk of postmenopausal breast cancer with higher adult weight gain among those with greater waist circumference. Despite a non-significant p for interaction, the point estimates are all elevated suggesting that there may be a nonlinear relationship and effect modification between them. No other studies have examined interactions between weight gain and central adiposity. It has been suggested that central adipose tissue is more metabolically active than peripheral adipose.[18] Our findings would suggest that only among women with greater central adipose is weight gain associated with risk. In addition, there has been speculation that timing of weight gain may have different effects on the location of fat disposition.[28,29] Further, greater central adiposity may reflect differences in steroid hormones including testosterone.[30,31] As seen in other studies of weight gain and breast cancer risk, findings from previous studies have generally shown increased risk of postmenopausal breast cancer in relation to various measures of central adiposity in most[7-9,19], but not in all studies.[26,27] Waist-hip ratio was the commonly used measure of

central adiposity, but some recent studies have used waist circumference which has been shown to be a stronger predictor of breast caner risk than waist-hip ratio.[9,18] These interactions between central adiposity and weight gain during the period of hormonal changes may indicate the importance of metabolic changes in relation to breast cancer risk. There is evidence linking central adiposity to altered glucose metabolism. It could also be that changes in glucose metabolism associated with increased central body fat explained the observed associations.

To further evaluate timing of weight gain on the risk, we examined effects of early life body-size indicators on breast cancer risk. Recent studies have shown some evidence that obesity in childhood is protective,[14,32] and higher growth rates during adolescence may be an independent risk factor of breast cancer in adulthood.[33,34] When we asked about a crude measure of body shape at the time of menarche, we did not find any association of this measure with risk. As have others, we found a protective effect of higher BMI at age 20; women in the highest quartile had a risk of 0.73 (95% CI 0.55-0.97), supporting hypothesis that a heavy build in early adulthood reduces the risk of breast cancer in adulthood. In addition, height, as a marker of exposure to nutritional factors in early life, was associated with increased postmenopausal breast cancer risk as we presented previously.

In our study, there were a small number of women who reported losing weight during their adult lives or during one of the time periods. For most measures, there were non-significant increases in risk in this group. We also looked at risk associated with weight gain when this group was included in the referent. Findings were similar to those presented with no weight gain as the referent. Further, because of the possibility of

weight loss in late stage breast cancer patients around the time of diagnosis, we examined results restricted to early stage breast cancer; findings were not different from those for all breast cancer cases shown here (data not shown). Additionally, because all models in our study were adjusted for the frequency matching factor of age but not for the other two factors, we confirmed that excluding those variables did not affect the point estimates.

Strengths of this population-based case-control study include the assessment of adult weight change at numerous time periods. We were able to assess various indicators of weight change and central adiposity, especially during periods of hormonal change in a woman's life, and to assess effect modification of this association by stratifying analyses by hormone receptor status and the use of hormone replacement therapy. We were also able to make multiple comparisons between time periods. However, it is possible that the use of different cut points for the categorical analyses may limit our ability to infer from the data.

There are several limitations which need to be considered in interpretation of these results. These include those common to the case-control study design: recall and selection bias as well as misclassification in the self-reported measures. In order to assess selection bias, we obtained general information from non-participants in a brief phone interview, and found that non-participants were not different from those participants for several characteristics, including diet. However, we did not query non-participants regarding anthropometry. There may be recall bias in the lifetime self-reported weight. We used the same methods for both cases and controls, and there was no particular emphasis on weight-related questions. Interviewers were blinded as to case and control status. In addition, reported weight one year before interview had a high correlation with

measured weight at the time of interview, and was similar for cases and controls (0.92 and 0.91 for cases and controls, respectively). For the lifetime weight change, we used reported weight for age 20 and for one year before interview. We were concerned that for cases the current measured weight might be affected by their disease or by treatment. In fact, there were few differences. For waist circumference we relied on the interviewer measurements.

This study supports epidemiologic evidence showing an increased risk of breast cancer associated with adult weight gain among postmenopausal women, and suggests that weight gain during the 30's and 40's, weight gain since a woman's first pregnancy, and weight gain since menopause may all be of importance in relation to postmenopausal breast cancer risk. Our results corroborate the evidence from previous studies, demonstrating interactions of weight gain and central adiposity measures, and suggesting that weight gain during periods of hormonal change are associated with higher risk. Further, the findings of an association with risk limited to ER+ or PR+ tumors and of stronger risk associated with never users of hormone replacement therapy also appear to provide evidence that the mechanism of an association of body weight with risk includes a hormonal etiology. Timing of adult weight gain appears to be of importance in relation to risk; evaluation of weight gain during times of hormonal changes and in relation to central adiposity should be evaluated in other research settings.

References

1. Hunter DJ, Willett WC. Diet, body size, and breast cancer. Epidemiol Rev 1993;15:110-32.
2. Cold S, Hansen S, Overvad K, Rose C. A woman's build and the risk of breast cancer. Eur J Cancer 1998; 34:1163-74.
3. International Agency for Research on Cancer (IARC). Handbooks of cancer prevention: weight control and physical activity. Lyon: IARC Press;2002.
4. Friedenreich C. Review of anthropometric factors and breast cancer risk. Eur J Cancer Prev 2001;10:15-32.
5. Van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, Fraser G, Goldbohm RA, Graham S, Kushi L, Marshall JR, Miller AB, et al. Pooled Analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol 2000;152:514-27.
6. Ursin, G. Longnecker, MP. Haile, RW. Greenland S. A meta-analysis of body mass index and risk of premenopausal breast cancer. Epidemiol 1995;6:137-41.
7. Ballard-Barbash R, Schatzkin A, Carter CL, Kannel WB, Kreger BE, D'Agostino RB, Splansky GL, Anderson KM, Helsel WE. Body fat distribution and breast cancer in the Framingham Study. J Natl Cancer Inst 1990;82: 286-90.
8. Folsom AR, Kaye SA, Prineas RJ, Potter JD, Gapstur SM, Wallace RB. Increased incidence of carcinoma of the breast associated with abdominal adiposity in postmenopausal women. Am J Epidemiol. 1990;131:794-803.
9. Huang Z, Willett WC, Colditz GA, Hunter DJ, Manson JE, Rosner B, Speizer FE, Hankinson SE. Waist circumference, waist hip ratio, and risk of breast cancer in the Nurses' Health Study. Am J Epidemiol 1999;150:1316-24.

10. Le Marchand L, Kolonel LN, Earle ME, Mi MP. Body size at different periods of life and breast cancer risk. Am J Epidemiol 1988;128:137-52.
11. Ballard-Barbash R, Schatzkin A, Taylor PR, Kahle LL. Association of change in body mass with breast cancer. Cancer Res 1990;50:2152-55.
12. Coates RJ, Uhler RJ, Hall HI, Potischman N, Brinton LA, Ballard-Barbash R, Gammon MD, Brogan DR, Daling JR, Malone KE, Schoenberg JB, Swanson CA. Risk of breast cancer in young women in relation to body size and weight gain in adolescence and early adulthood. Br J Cancer 1999;81:167-74.
13. Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, Manson JE, Hennekens CH, Rosner B, Speizer FE, Willett WC. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997;278:1407-11.
14. Magnusson C, Baron J, Persson I, Wolk A, Bergstrom R, Trichopoulos D, Adami HO. Body size in different periods of life and breast cancer risk in post-menopausal women. Int J Cancer 1998;76:29-34.
15. Trentham-Dietz A, Newcomb PA, Egan KM, Titus-Ernstoff L, Baron JA, Storer BE, Stampfer M, Willett WC. Weight change and risk of postmenopausal breast cancer (United States). Cancer Causes Control 2000;11:533-42.
16. Stoll BA. Perimenopausal weight gain and progression of breast cancer precursors Cancer Detect Prev 1999;23:31-36.
17. Harvie M, Howell A, Vierkant RA, Kumar N, Cerhan JR, Kelemen LE, Folsom AR, Sellers TA. Association of gain and loss of weight before and after menopause with risk of postmenopausal breast cancer in the Iowa Women's Health study. Cancer Epidemiol Biomarkers Prev 2005 14:656-61.

18. Ballard-Barbash R. Anthropometry and breast cancer. body size-a moving target. Cancer 1994;74:1090-100.
19. Kaaks R. Van Noord PA. Den Tonkelaar I. Peeters PH. Riboli E. Grobbee DE. Breast-cancer incidence in relation to height, weight and body-fat distribution in the Dutch DOM cohort. Int J Cancer 1998;76:647-51.
20. Willett WC. Nutritional Epidemiology. Oxford:Oxford University Press;1998.
21. Bonner MR. Han D. Nie J. Rogerson PA. Vena JE. Muti P. Trevisan M. Edge S. Freudenheim JL. Breast caner risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. Cancer Epidemiol Biomarkers Prev 2005;14:53-60.
22. Han D. Rogerson PA. Bonner MR. Nie J.Vena JE. Muti P. Trevisan M. Freudenheim JL. Assessing spatio-temporal variability of risk surfaces using residential history data in a case control study of breast cancer. Int J Health Geogr 2005;4:9.
23. Feigelson HS. Jonas CR. Teras LR. Thun MJ. Calle EE. Weight gain, body mass index, hormone replacement therapy, and postmenopausal breast cancer in a large prospective study. Cancer Epidemiol Biomarkers Prev 2004;13:220-24.
24. Friedenreich CM. Courneya KS. Bryant HE. Case-control study of anthropometric measures and breast cancer risk. Int J Cancer 2002;99:445-52.
25. Enger SM, Ross RK, Paganini-Hill A, Carpenter CL, Bernstein L. Body size, physical activity, and breast cancer hormone receptor status: results from two case-control studies. Cancer Epidemiol Biomarkers Prev 2000;9:681-87.

26. Lahmann PH, Hoffmann K, Allen N, van Gils CH, Khaw KT, Tehard B, Berrino F, Tjonneland A, Bigaard J, Olsen A, Overvad K, Clavel-Chapelon F, et al. Body size and breast cancer risk: findings from the European prospective investigation into cancer and nutrition (EPIC). Int J Cancer 2004;111:762-71.
27. Morimoto LM. White E. Chen Z. Chlebowski RT. Hays J. Kuller L, Lopez AM, Manson J, Margolis KL, Muti PC, Stefanick ML, McTiernan A. Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative (United States). Cancer Causes Control 2002;13:741-51.
28. Kumar NB. Lyman GH. Allen K. Cox CE. Schapira DV. Timing of weight gain and breast cancer risk. Cancer 1995;76:243-49.
29. Stoll BA. Timing of weight gain in relation to breast cancer risk. Ann Oncol 1995;6:245-8.
30. Potischman N, Swanson CA, Siiteri P, Hoover RN. Reversal of relation between body mass and endogenous estrogen concentrations with menopausal status. J Natl Cancer Inst 1996;88:756-58.
31. Stoll BA. Adiposity as a risk determinant for postmenopausal breast cancer. Int J Obes Relat Metab Disord. 2000;24:527-33.
32. Barnes-Josiah D. Potter JD. Sellers TA. Himes JH. Early body size and subsequent weight gain as predictors of breast cancer incidence (Iowa, United States). Cancer Causes Control 1995;6:112-18.
33. Ahlgren M, Melbye M, Wohlfahrt J, Sorensen T. Growth patterns and the risk of breast cancer in women. N Engl J Med 2004;351:1619-26.

34. Berkey C. Frazier A. Gardner J. Colditz G. Adolescence and breast carcinoma risk. Cancer 1999;85:2400-9.

Table 1. Descriptive characteristics of study participants (Mean ± Standard Deviation): WEB Study, 1996-2001.

	Premenopausal women (n=935)			Postmenopausal women (n=2336)		
	Cases (*n*=325)	Controls (*n*=610)	*p*-value	Cases (*n*=841)	Controls (*n*=1495)	*p*-value
Age (years)	44.9±4.6	44.1±4.6	.01	63.0±8.5	63.4±8.9	.30
Education (years)	14.0±2.3	14.2±2.2	.08	13.3±2.6	13.0±2.3	<.0001
Age at menarche (years)	12.5±1.6	12.6±1.6	.76	12.6±1.6	12.8±1.7	.02
Age at first pregnancy* (years)	25.0±5.1	25.8±4.8	.03	23.8±4.7	23.5±4.3	.16
Age at menopause (years)	-	-	-	48.3±5.4	47.4±6.3	.001
Time since first pregnancy* (years)	19.9±7.2	18.4±6.8	.01	39.5±9.2	40.1±9.3	.19
Time between first pregnancy and menopause* (years)	-	-	-	24.4±6.8	24.1±7.0	.31
Height (cm)	164.0±6.4	163.8±6.3	.63	161.8±6.6	160.8±6.1	<.0001
BMI one year ago (kg/m^2)	26.7±6.6	27.2±6.8	.28	28.9±6.1	28.4±6.4	.07
Weight at interview (kg)	73.1±19.2	73.9±19.0	.52	75.7±16.1	73.8±16.3	.01
Weight one year ago (kg)	71.7±18.4	72.9±19.2	.35	75.6±16.4	73.4±17.0	.003
Abdominal height (cm)	19.6±3.3	19.6±3.6	.94	21.3±3.2	20.9±3.3	.01
Waist circumference (cm)	84.4±15.0	84.2±15.2	.86	91.3±14.2	88.7±14.0	<.0001
BMI at age 20 (kg/m^2)	20.9±3.4	21.5±4.0	.03	21.0±3.0	21.4±3.2	.01
Weight at age 20 (kg)	56.3±10.2	57.6±11.2	.09	55.0±8.5	55.2±8.6	.59
Weight gained during first pregnancy*	15.9±8.6	14.9±7.8	.16	15.5±10.8	15.1±10.3	.35
Weight gain since age 20 (kg)	15.4±14.9	15.4±15.8	.35	20.6±14.6	18.2±15.4	<.0001
Body type at menarche[1]	2.2±1.2	2.4±1.4	.04	2.1±1.2	2.1±1.2	.55
Relative with breast cancer (yes)	21%	10%	<.0001	20%	14%	<.0001
Use of hormone replacement therapy* (yes)	-	-	-	54%	50.0%	.09

* Sample sizes vary for this group because of exclusion of missing data and those women with no pregnancy.

[1] One being thinest and nine being fattest

Table 2. Postmenopausal breast cancer risk associated with weight change during several key time periods with and without adjustment for BMI one year ago and for adult lifetime weight gain, WEB Study, 1996-2001.

Weight change: age 20 to one year ago (kg)	Cases (n=841)	Controls (n=1495)	Adjusted OR[1] (95%CI)	Adjusted OR[2] (95%CI)
≤ 0	47	131	0.84(.53-1.33)	0.90(.56-1.45)
0-9.1	137	330	1.00	1.00
9.1-17.7	208	343	1.49(1.10-2.01)	1.45(1.06-1.96)
17.7-27.3	227	360	1.61(1.19-2.17)	1.53(1.12-2.08)
>27.3	222	331	1.86(1.37-2.52)	1.71(1.23-2.37)
p-Trend			.001	.05
Weight change: age 20 to 29 (kg)	Cases (n=841)	Controls (n=1495)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	232	414	1.12(.82-1.53)	1.25(.92-1.71)
0-2.3	177	330	1.00	1.00
2.3-4.5	168	323	1.12(.82-1.53)	1.06(.78-1.45)
4.5-6.8	98	165	1.21(.83-1.77)	1.10(.75-1.61)
>6.8	166	263	1.40(.94-2.09)	1.08(.71-1.67)
p-Trend			.02	.07
Weight change: age 30 to 39 (kg)	Cases (n=841)	Controls (n=1495)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	225	456	0.89(.66-1.19)	0.95(.71-1.27)
0-2.3	179	346	1.00	1.00
2.3-4.5	181	313	1.19(.87-1.61)	1.11(.82-1.52)
4.5-6.8	113	151	1.52(1.06-2.18)	1.35(.93-1.96)
>6.8	143	229	1.72(1.14-2.59)	1.32(.83-2.09)
p-Trend			.05	.04
Weight change: age 40 to49 (kg)	Cases (n=840)	Controls (n=1494)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	214	468	0.86(.56-1.31)	0.87(.57-1.32)
0-2.3	60	112	1.00	1.00
2.3-4.5	155	327	0.99(.65-1.53)	0.96(.63-1.48)
4.5-7.8	248	331	1.71(1.13-2.59)	1.58(1.02-2.43)
>7.8	163	256	1.92(1.17-3.15)	1.60(.91-2.79)
p-Trend			.05	.07
Weight change: age 50 to 59 (kg)	Cases (n=790)	Controls (n=1357)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	218	441	0.99(.73-1.34)	1.05(.77-1.43)
0-2.3	153	278	1.00	1.00
2.3-4.5	182	253	1.52(1.11-2.09)	1.39(1.01-1.93)
4.5-9.1	70	145	0.88(.58-1.33)	0.76(.50-1.17)
>9.1	167	240	1.32(.87-1.99)	0.91(.56-1.48)
p-Trend			.05	.26
Weight change: age 60 to 69 (kg)	Cases (n=497)	Controls (n=925)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	170	331	0.99(.96-1.02)	1.09(.75-1.58)
0-2.3	92	214	1.00	1.00

2.3-4.5	82	131	1.33(.88-2.03)	1.19(.78-1.83)
4.5-8.3	44	101	1.06(.64-1.76)	0.90(.53-1.52)
>8.3	109	148	1.94(1.21-3.10)	1.27(.72-2.27)
p-Trend			.02	.06
Weight change: age first pregnancy to one year ago (kg)	Cases (n=693)	Controls (n=1340)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	43	150	0.75(.49-1.16)	0.74(.48-1.16)
0-7.7	112	289	1.00	1.00
7.7-14.1	175	303	1.48(1.10-2.01)	1.50(1.09-2.05)
14.1-22.7	174	303	1.53(1.12-2.09)	1.57(1.09-2.25)
>22.7	189	295	1.70(1.22-2.37)	1.78(1.08-2.94)
p-Trend			.06	.36
Weight change: age menopause to one year ago (kg)	Cases (n=841)	Controls (n=1495)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	203	436	0.87(.64-1.19)	0.91(.66-1.25)
0-3.6	141	254	1.00	1.00
3.6-7.3	163	272	1.03(.74-1.43)	0.99(.71-1.38)
7.3-13.6	133	243	1.22(.86-1.73)	1.11(.76-1.61)
>13.6	201	290	1.58(1.09-2.30)	1.27(.78-2.05)
p-Trend			.13	.36
Weight change: age first pregnancy to menopause (kg)	Cases (n=693)	Controls (n=1340)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	96	229	1.15(.80-1.66)	1.22(.84-1.76)
0-4.5	91	243	1.00	1.00
4.5-8.2	166	309	1.58(1.14-2.20)	1.48(1.06-2.06)
8.2-13.6	190	297	1.90(1.36-2.66)	1.63(1.13-2.33)
>13.6	150	262	1.91(1.26-2.88)	1.34(.81-2.22)
p-Trend			.07	.01
Weight change: age 20 to age first pregnancy (kg)	Cases (n=693)	Controls (n=1340)	Adjusted OR[2] (95%CI)	Adjusted OR[3] (95%CI)
≤ 0	244	492	0.99(.73-1.35)	1.03(.76-1.41)
0-2.3	117	236	1.00	1.00
2.3-4.5	117	258	0.98(.70-1.36)	0.94(.67-1.32)
4.5-7.6	76	142	1.15(.78-1.72)	1.08(.72-1.61)
>7.6	139	212	1.52(.98-2.36)	1.25(.79-1.98)
p-Trend			.02	.07

[1] Adjusted for age, education, previous benign breast disease, age at menarche, age at first birth, family history of breast cancer, and age at menopause, use of hormone replacement therapy.
[2] Adjusted for age, education, previous benign breast disease, age at menarche, age at first birth, family history of breast cancer, and age at menopause, use of hormone replacement therapy, BMI residuals.
[3] Adjusted for age, education, previous benign breast disease, age at menarche, age at first birth, family history of breast cancer, and age at menopause, use of hormone replacement therapy, BMI residuals, weight change between age 20 and one year before interview.

Table 3. Postmenopausal breast cancer risk associated with weight change between age 20 and one year ago, stratified by waist circumference

Weight change (kg)	Below median (<88cm)			Above median (≥88 cm)		
	Cases (n=366)	Controls (n=745)	Adjusted OR* (95%CI)	Cases (n=475)	Controls (n=750)	Adjusted OR* (95%CI)
≤ 0	38	105	.86(.50-1.48)	9	26	1.61(.49-5.20)
0-9.1	123	277	1.0	14	53	1.0
9.1-17.7	126	215	1.22(.84-1.76)	82	128	2.48(1.16-5.28)
17.7-27.3	67	131	1.10(.68-1.76)	160	229	2.39(1.15-4.95)
>27.3	12	17	1.31(.52-3.27)	210	314	2.33(1.11-4.90)
p-Trend			.44			.69

* Adjusted for age, education, age at menarche, age at first birth, previous benign breast disease, family history of breast cancer, age at menopause, BMI residuals, HRT use.

Table 4. Postmenopausal breast cancer risk associated with weight change between age 20 and one year ago, stratified by estrogen and progesterone receptor status and by HRT use

Weight change (kg)	ER+			ER-		
	Cases (n=510)	Controls (n=1495)	Adjusted OR[1] (95%CI)	Cases (n=136)	Controls (n=1495)	Adjusted OR[1] (95%CI)
≤ 0	30	131	1.03(.58-1.82)	10	131	1.31(.51-3.33)
0-9.1	69	330	1.0	24	330	1.0
9.1-17.7	124	343	1.60(1.09-2.35)	30	343	1.48(.78-2.80)
17.7-27.3	138	360	1.86(1.27-2.73)	42	360	1.62(.86-3.05)
>27.3	149	331	2.42(1.62-3.61)	30	331	1.19(.58-2.43)
p-Trend			.001		.68	
Weight change (kg)	PR+			PR-		
	Cases (n=389)	Controls (n=1495)	Adjusted OR[1] (95%CI)	Cases (n=257)	Controls (n=1495)	Adjusted OR[1] (95%CI)
≤ 0	20	131	1.05(.52-2.12)	20	131	1.13(.58-2.18)
0-9.1	45	330	1.0	48	330	1.0
9.1-17.7	94	343	2.07(1.32-3.27)	60	343	1.14(.72-1.82)
17.7-27.3	107	360	2.39(1.52-3.76)	73	360	1.33(.84-2.11)
>27.3	123	331	3.14(1.96-5.04)	56	331	1.21(.73-2.01)
p-Trend			.001		.82	
Weight change (kg)	Never-users of HRT			Ever-users of HRT		
	Cases (n=385)	Controls (n=716)	Adjusted OR[2] (95%CI)	Cases (n=447)	Controls (n=717)	Adjusted OR[2] (95%CI)
≤ 0	20	64	0.75(.34-1.63)	27	62	1.03(.56-1.89)
0-9.1	49	146	1.00	88	174	1.00
9.1-17.7	96	160	1.69(1.04-2.75)	109	171	1.32(.88-1.97)
17.7-27.3	97	164	1.86(1.15-3.03)	127	182	1.38(.91-2.08)
>27.3	123	182	2.00(1.22-3.27)	96	128	1.52(.96-2.41)
p-Trend			.12			.26

[1] Adjusted for age, education, age at menarche, age at first birth, previous benign breast disease, family history of breast cancer, age at menopause, BMI residuals, HRT use.

[2] Adjusted for age, education, age at menarche, age at first birth, previous benign breast disease, family history of breast cancer, age at menopause, BMI residuals.